# MSURJ

MCGILL SCIENCE UNDERGRADUATE RESEARCH JOURNAL

VOLUME 21 | ISSUE 1| APRIL 2026

# ON THE COVER

What is required for a seedling to grow? Perhaps water, soil, sunlight, and even electrical stimulation!

MSURJ's 21st volume cover art reflects the potential benefits of electrical stimulation on plants to enhance their growth. Inspired by author Bryan Eli Khoury's research article published in this volume, the cover art piece offers a vibrant, fluid, and luscious depiction of the effects of electrical input on germination and seedling development.

While electrical stimulation was applied in Khoury's research work through wiring and alligator clips attached to corn seeds, artist Auden Akinc reimagines this process by depicting luminescent currents traveling through winding conduits, electrifying emerging plant shoots. The glowing pathways suggest the influence of electrical input on plant development, particularly in modulating shoot growth. Subtle sparks radiating across the leaves evoke this potential. Created in a digital medium, this cover presents an organic abstraction of the interplay between electrical stimulation and plant development.

Auden Akinc is a second year McGill student majoring in physics and minoring in biology. In high school, he majored in Visual Arts at Walnut Hill School for the Arts, working primarily on oil painting, ceramic sculpture, apparel design and figure illustration. He is passionate about astrophysics, ecology, and the arts.

MCGILL SCIENCE
UNDERGRADUATE
RESEARCH JOURNAL

# BOARD 2025-2026




Email: msurj.sus@mail.mcgill.ca
Website: https://msurjonline.mcgill.ca/about/editorialTeam

# TABLE OF CONTENTS

# FOREWORD

**Dear Reader,**

**The twenty-first volume of the *McGill Science Undergraduate Research Journal* comes on the heels of celebrating two decades of publishing.** Last year, we took the opportunity to look back at how far the journal and, by extension, undergraduate research has come. If Volume 20 was about where we had been, then Volume 21 is decidedly about where we are going.

Milestones deceive. They invite us to view our progress as a project that has been completed rather than as a continual process. This year has marked a turning point in science globally. The enterprise of science is changing rapidly and sustaining a research journal within it requires work renewed each year by a new board, new authors, and new questions. In so doing, each year, we earn the platform we have been given.

This volume features nine articles, including MSURJ's first-ever Commentary, reflecting the breadth of undergraduate investigation. In particular, we are proud to publish work spanning scales: from using electricity to stimulate germination in plants, to characterizing crayfish species composition in the Gault Nature Reserve, to modelling our peatland carbon sinks from space. Undergraduate research remains unconfined by scope, discipline, or scale.

What these works share is an orientation: each author chose to subject their findings to the scrutiny of peer review, to revise and defend their conclusions. That process, as much as the published result, is what our journal exists to foster.

We present to you *Volume 21* with immense pride in our past and excitement for the future of undergraduate scientific inquiry.

On behalf of our entire Editorial Board, thank you.

**Lisa Xie and Benjamin Lévesque Kinder**
**Editors-in-Chief**

# ACKNOWLEDGEMENTS

The *McGill Science Undergraduate Research Journal* would like to thank its generous contributors, without whom this journal would not have been possible.

We thank the librarians and staff of McGill University for their guidance, especially Jennifer Innes for the time and technical support she has given us. We thank the Office of Science Education for their continued support and collaboration. Furthermore, we thank our advisors Dr. Véronique Brulé, Dr. Christie D. Rowe, and John Ni for their counseling and encouragement.

We thank all our financial supporters in the McGill community for their generous support:

Science Undergraduate Society of McGill University
Faculty of Science
Faculty of Medicine and Health Sciences, School of
Biomedical Sciences
Department of Physiology
Centre for Structural Biology Research

Finally, we would like to thank all of our anonymous peer reviewers. Their careful revision and feedback is critical to the success of this volume.

Evelyn Huaman[1], Pierre Demerjian [2]

# Engineering Safety in Childhood Bone Fragility: Built Environments for Fracture Prevention

Commentary [1]Ingram School of Nursing, McGill University; [2]Department of Computer Science, McGill University **Correspondence:** evelyn.huaman2@mail.mcgill.ca

Submitted: 21/02/2026 | Accepted: 08/03/2026 | Published: 4/10/2026

https://doi.org/10.26443/msurj.v21i1.355 

Childhood bone fragility conditions affecting balance and coordination share a common feature: fractures often follow minor impacts in ordinary spaces.[1] In these disorders, risk is shaped not only by bone biology, but also by modifiable features of the built environment.[2] Designing environments that reduce these impact hazards for children with bone fragility ultimately makes everyday spaces safer and more accessible for everyone. This issue is exemplified in Osteogenesis imperfecta (OI), a heritable connective tissue disorder characterized by bone fragility and fractures after minimal trauma. OI is often accompanied by skeletal deformity that threatens mobility.[1] Almost half of children and one-third of adults with OI experience a fracture within one year of diagnosis.[2] Fracture risk in pediatric OI is informed by exposure to environmental hazards, with the built environment directly influencing injury frequency and severity. OI, therefore, requires lifelong hazard circumvention, which often leads to exhaustion and isolation.[3]

Current OI management combines medical and rehabilitative interventions, including bone-active therapy and corrective surgery.[4] This approach treats OI-related disability as a biological and behavioural issue and neglects the environmental contributions to fracture and disability risk.[3,5]

Fracture risk arises from the combined effects of the interactions between bone fragility, hazardous exposures, and the mechanics of everyday injury. In other words, fractures are not only the result of weak bones but also of cumulative exposure to hazards in everyday environments. These risks are embedded in everyday environments, from the bathroom and the school corridor to playgrounds designed around speed, height and impact. These environmental hazards increase the likelihood of fractures, even when bone health and clinical care are optimized.[6]

We propose that the environmental design of childrens spaces require planning and resource allocation as part of OI care, comparable to pharmacological and surgical interventions. While we focus here on pediatric OI, similar environmental interventions could meaningfully reduce fracture and injury risk in children with conditions such as cerebral palsy, Duchenne muscular dystrophy, or ataxic syndromes. Furthermore, geriatric populations that are also at high risk for fractures, and in which there is a high prevalence of osteoporosis, Parkinsons disease, and other conditions, also benefit from these interventions.

The intentional design of spaces can predictably alter the frequency of hazard exposure and the likelihood of injuries. In safety science, the hierarchy of controls ranks methods of hazard reduction, from most to least effective: hazard elimination, safer substitutions, spatial redesign, rules and supervision, and, as a last line of defence, protective gear.[7] This perspective shifts fracture prevention from relying on individual vigilance toward systematic environmental design that proactively reduces fall risk and limits impact hazards. Moreover, the International Classification of Functioning, Disability and Health by the World Health Organization explicitly recognizes environmental factors as determinants of individual functioning, participation, and health.[8] Thus, OI management that ignores built environments remains incomplete.[9]

> Children often lose access to play and peer participation, not because safer alternatives are impossible, but because systems default to restriction rather than redesign.

Across settings, small design decisions and routine policies are modifiable risk factors. At home, steep staircases without railings, slippery floors, and cluttered pathways increase fall risk; simple home modifications that make transfers and movement safer, such as grab bars, slip-resistant surfaces, and adequate space around key fixtures can reduce everyday hazards.[10] School environments can pose their own barriers, from poor circulation and inaccessible exits to playgrounds built around impact. Conversely, accessible classrooms and play spaces, with step-free routes and equipment that more children can use safely, support mobility and social inclusion.[11] In recreational settings, if most community sports are high-impact activities, many children with OI will find themselves with limited ways to participate, restricting their opportunities to move confidently and remain part of shared play.[9,12]

When spaces are not inclusive, the burden of mitigating fracture risk falls on the child and family, who are often advised to exercise caution by limiting high-impact activities and using mobility aids.[5] These measures, in turn, can lead to isolation.[13] Children often lose access to play and peer participation, not because safer alternatives are impossible, but because systems default to restriction rather than redesign. This tendency is rooted as much in liability concerns, as in constrained resources that limits investment in accessible, inclusive environments.[14] This same logic applies to seniors and individuals with disabilities. An equitable approach should mitigate risks by reducing hazards while preserving the individuals right to participate in school and community life.[15] Crucially, supervised and personalized activity programs can improve physical capacity and everyday function in children with OI.[3]

Implementation of inclusion strategies depends as much on policy and culture as on architecture. Unfortunately, attention to inclusive environmental design remains limited by cost, scarce OI-specific environmental research, and responsibility spread too thinly across sectors.[11,16,17] Furthermore, clinicians, educators, and community providers may overestimate injury risk for children with disabilities, and default to overly cautious restrictions rather than to personalized care focused on creating opportunities.[11,15] Schools and municipalities also face competing demands that delay accessibility investments, reinforcing a pattern in which risk is managed through exclusion rather than redesign.[11] Addressing these barriers requires coordinated policy, research, and design efforts that embed inclusion and safety in everyday environments.

Guiding families through identifying modifiable hazards and holistic environmental adjustments ensures that clinical care follows families home. Tracking OI patients physical activity and near-fall events, alongside fractures, would also improve the ability of health providers to advise families on common hazards and on strategies to mitigate risk in these areas. Universal design, which plans spaces to accommodate diverse needs from the outset, makes environments easier to move through and use safely. From step-free entry to clearer wayfinding, universal design maximizes accessibility across various populations. Adopting universal design at community-level institutions, such as schools and recreation centers, ensures that the responsibility for minimizing predictable hazards is institutionalized and does not rely on individual families advocacy. The implementation and prioritization of sustainable adaptive programming through comprehensive staff training and investment in accessible equipment also removes barriers to participation for families with OI diagnoses. Increased safety and risk mitigation alongside adaptive programming improves safe participation for children with OI and distributes the responsibility of care more evenly amongst the people around them who navigate the same environments.[18] Similarly to other accessibility requirements, public health and education systems should take on the responsibility of managing structured environmental audits to ensure adequate hazard reduction.[11] This approach has shown promising results in other contexts. For instance, when schools replaced unsafe playground equipment with equipment that met safety standards, playground injury rates decreased significantly.[19]

While environmental factors are widely recognized as important components of fracture risk, direct evidence linking specific environmental modifications to reduced fracture rates in OI remains limited. Future research should evaluate environmental interventions for their effects on injury risk, mobility, and day-to-day well-being; assess implementation performance, including whether these interventions can actually be afforded, adapted, and sustained in real settings over time; and develop OI-specific environmental guidance. Positioning environments alongside medical care aligns OI management with both medical and social models of disability, underscoring how non-clinical interventions spanning architecture, programming, and equipment can shape day-to-day functioning and offer a potential path to improved physical and mental well-being for children living with OI. What OI makes visible is a reality that applies to all bodies: environments shape health outcomes, making the case for accessible design not a niche concern but a universal one.

## References


1. Ward, L. M. A practical guide to the diagnosis and management of osteoporosis in childhood and adolescence. Front. Endocrinol. 14 (2024).
2. Iyer, S., Hughes Garza, H. & Lawson, K. A. Femur fractures in children under two related to consumer products and the home environment treated in United States emergency departments 2017–2021. Inj. Prev. 32, 163–167 (2024).
3. Al Arab, H. et al. Evaluation of the benefits of adapted physical activity in children and adolescents with osteogenesis imperfecta: the MOVE-OI trial. Orphanet J. Rare Dis. 20, 175 (2025).
4. Nijhuis, W., Verhoef, M., Van Bergen, C., Weinans, H. & Sakkers, R. Fractures in osteogenesis imperfecta: pathogenesis, treatment, rehabilitation and prevention. Children 9, 268 (2022).
5. Montpetit, K., Palomo, T., Glorieux, F. H., Fassier, F. & Rauch, F. Multidisciplinary treatment of severe osteogenesis imperfecta: functional outcomes at skeletal maturity. Arch. Phys. Med. Rehabil. 96, 1834–1839 (2015).
6. World Health Organization. World report on child injury prevention 2008. `https://www.who.int/publications/i/item/9789241563574`.
7. Centers for Disease Control and Prevention. About hierarchy of controls 2024. `https://www.cdc.gov/niosh/hierarchy-of-controls/about/index.html`.
8. World Health Organization. International classification of functioning, disability and health (ICF) 2024. `https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health`.
9. Labanca, L. et al. Clinical-functional features of individuals with osteogenesis imperfecta and Ehlers–Danlos syndromes: a scoping review of assessment tools and ICF model. Musculoskelet. Sci. Pract. 64, 102742 (2023).
10. Marr, C., Seasman, A. & Bishop, N. Managing the patient with osteogenesis imperfecta: a multidisciplinary approach. J. Multidiscip. Healthc. 10, 145–155 (2017).
11. Cleary, B. et al. The school of no fun and no play: learning how to create supportive environments from children with osteogenesis imperfecta. J. Disabil. Stud. Educ. 4, 25–49 (2024).
12. Gilani, M. et al. Acceptability and barriers of exercise in children with osteogenesis imperfecta. Arch. Rehabil. Res. Clin. Transl. 7, 100458 (2025).
13. Galhardo Saraiva, F. et al. Does physical activity play a role in the management of children and young adults with osteogenesis imperfecta? Cureus (2024). https://doi.org/10.7759/cureus.53646
14. Jerebine, A., Fitton-Davies, K., Lander, N., et al. Children are precious cargo; we dont let them take any risks!: Hearing from adults on safety and risk in childrens active play in schools. Int. J. Behav. Nutr. Phys. Act. 19, 111 (2022). https://doi.org/10.1186/s12966-022-01344-7
15. Carbone, P. S., Smith, P. J., Lewis, C. & LeBlanc, C. Promoting the participation of children and adolescents with disabilities in sports, recreation, and physical activity. Pediatrics 148, e2021054664 (2021).
16. Rapoport, M. et al. The patient clinical journey and socioeconomic impact of osteogenesis imperfecta: a systematic scoping review. Orphanet J. Rare Dis. 18, 34 (2023).
17. Castro, A. R. et al. The day-to-day experiences of caring for children with osteogenesis imperfecta: a qualitative descriptive study. J. Clin. Nurs. 29, 2999–3011 (2020).
18. Moore, A., Lynch, H. & Boyle, B. Can universal design support outdoor play, social participation, and inclusion in public playgrounds? A scoping review. Disabil. Rehabil. 44, 3304–3325 (2020).
19. Obrusnikova, I., Firkin, C. J., Pennington, R., Dixon, I. & Bilbrough, C. Statewide assessment of public park accessibility and usability and playground safety. Int. J. Environ. Res. Public Health 23, 139 (2026).

Submitted: 01/08/2026
Accepted: 02/19/2026
Published: 04/10/2026

Research Article

[1]Faculty of Science, McGillUniversity, Montréal, QC, Canada

Keywords

Zea mays, electrical stimulation, germination, voltage, growth rate

Email Correspondence

bryan-eli.khoury@mail.mcgill.ca

https://doi.org/10.26443/msurj.v21i1.391



Khoury, Bryan-Eli[1]

# Effects of Electrical Stimulation on Germination and Early Seedling Growth in Corn

## Abstract

Early seedling establishment is increasingly important for agricultural resilience[1]. Electrical stimulation has been proposed as a low-cost approach to modulate plant signalling pathways and potentially improve germination and early growth[2]. This study tested whether brief daily electrical stimulation affects germination success and early seedling growth in corn (*Zea mays* subsp. *mays)*.

Pre-soaked corn seeds were distributed into five groups (0 V, 3 V, 6 V, 9 V, 12 V; 35 seeds per group; total *n*=175). Electrical stimulation was applied for 10 min daily over 5 consecutive days using battery power sources and alligator clip electrodes placed diagonally across each tray on a moistened paper substrate. Germination was defined by visible root emergence (with or without shoot emergence). After 5 days, shoot and root lengths were measured, and group means were compared using one-way ANOVA.

Germination percentages were similar across treatments (45.7-54.3%), with no clear voltage-dependent trend. Mean shoot length was highest at 6 V (1.20 cm; standard deviation (SD)=0.97) but did not differ significantly across voltages (ANOVA $p$=0.410). Mean root length peaked at 6 V (2.00 cm; SD=1.34) and decreased at higher voltages (e.g., 12 V: 1.10 cm; SD=0.85), with voltage producing a significant effect on root length (ANOVA $p$=0.014).

Interpretation is limited by the short duration (5 days) and substantial non-germination, which reduced the effective sample size and statistical power. Growth on a nutrient-free paper substrate may limit generalizing to soil conditions. In addition, results may be influenced by variability in delivered voltage (battery drift/electrode placement) and by measurement error from bent or branching roots.

Electrical stimulation did not improve germination, but moderate stimulation (6 V) was associated with enhanced early root elongation, while shoot growth differences were not significant. These findings suggest that appropriately tuned electrical stimulation may selectively enhance early root development in maize seedlings, warranting longer-term studies assessing downstream effects on plant performance and yield.

## Introduction

The challenge of enhancing seed germination and early seedling growth is critical in the field of agriculture, especially given the global issue of climate change, which poses significant challenges to seed germination and plant life[3]. Because establishment depends on multiple interacting conditions, a myriad of factors can affect germination and growth rates, such as temperature, water availability, light exposure, and electric stimulation. Therefore, this investigation aims to assess the influence of varying voltages of electrical stimulation (0 V, 3 V, 6 V, 9 V, 12 V) on the germination and growth of corn (*Zea mays* subsp. *mays)* seeds, with a focus on measuring changes in shoot and root length. Specifically, the study evaluates whether an intermediate voltage enhances early growth relative to both no stimulation and higher-voltage exposure.

Seed germination is a crucial stage in a plant's life cycle that follows a period of dormancy and is influenced by many internal and external factors. Because germination is highly sensitive to abiotic conditions, other factors such as water availability, oxygen exposure, and temperature were kept consistent across treatments. A dormant seed contains approximately 6-15% water in its cells, whereas metabolically active cells require about 75-95% water to sustain enzymatic and respiratory activity[4]. Therefore, the absorption of water by the embryo triggers cell rehydration and reactivates the seed's metabolism[5]. This rehydration also softens the seed coat, enhances permeability, and transforms food from an insoluble form to a soluble one while facilitating gas exchange. As metabolism resumes, adequate oxygen is required, as it is crucial for aerobic respiration to release the energy needed for early growth. Because respiration and other enzyme-mediated metabolic processes are temperature-sensitive, temperature also plays an important role in germination. Optimal temperatures around 30°C facilitate efficient enzymatic processes for most seeds (including *Zea mays* subsp. *mays)*, while cooler temperatures slow these reactions, and higher temperatures can denature enzymes, thus affecting and even preventing germination from occurring[6, 7].

Electrical signalling naturally occurs internally in both plants and animals. Similar to internal signalling, external electrical stimulation can regulate many plant biological processes at the cellular level by modifying cellular metabolism through interaction with chemical, hydraulic, and hormone signalling; at the biochemical level by controlling protein synthesis, secondary metabolites, hormones (jasmonic acid and abscisic acid), and protein kinase inhibitors; at the molecular level by regulating gene expression

of enzymes related to ribosome proteins and carbon metabolism; and at the physiological level by adjusting stem growth, leaf movement, stomatal opening, transpiration, and respiration[8]. In addition, electrical stimulation can significantly improve the plant's tolerance to freezing and salt stress[9].

Electrical stimulation can enhance seed germination by first activating ion channels, which then trigger $Ca^{2+}$ and ROS (reactive oxygen species) signalling. This signalling cascade increases the activities of catalase, peroxidase, and other enzymes related to carbon metabolism, and can enhance the uptake of nutrients like potassium (K), calcium (Ca), magnesium (Mg), etc. Beyond germination, external electrical stimulation may influence early seedling growth and post-germination physiology by promoting photosynthetic thermotolerance, defined here as the ability of the photosynthetic apparatus (particularly photosystem I) to maintain function under heat stress[9]. Following electrical stimulation, signalling molecules such as $Ca^{ref2+}$, ROS, and plant hormones can be mobilized, with changes in gene expression (e.g., calmodulin, proteinase inhibitors, and Rubisco) and phloem mass flow. Downstream, this response has been linked to decreased carbon dioxide ($CO_2$) assimilation and activation of photoprotective mechanisms, including non-photochemical quenching and cyclic electron flow, increased ATP content, and changes in transpiration, leaf heating, and state transitions. However, high-voltage electrical stimulation can inhibit seed germination by disrupting cell membrane integrity, which can impair water balance and lead to reduced relative water content, i.e., dehydration. This dehydration-associated stress can trigger a broad stress response, including the accumulation of secondary metabolites, and can increase oxidative damage; consequently, lipid peroxidation rises and malondialdehyde (MDA) accumulates as an indicator of stress-related cellular damage[10]. Antioxidant enzyme activity can decrease under high-voltage electrical stimulation, promoting $H_2O_2$ buildup and oxidative stress[11]. In shoots, a reduction in Superoxide dismutase (SOD) activity occurs, an important antioxidant enzyme that catalyzes the disproportionation of the superoxide anion to hydrogen peroxide, which can allow reactive oxygen species to accumulate and damage cellular components. However, catalase activity decreases, reducing $H_2O_2$ detoxification and increasing oxidative damage to membranes, proteins, and DNA. In roots, a decrease in ascorbate peroxidase (APX) activity occurs, further limiting hydrogen peroxide detoxification and antioxidant defence[9, 12].

While the concept of using electricity for seed germination is a promising and low-cost approach, it remains relatively underutilized and uninvestigated in the field of agricultural research. This paper aims to expand our understanding of how external electrical stimulation interacts with seeds to promote their germination and growth. Because electrical stimulation can influence ion-channel activity, $Ca^{2+}$/ROS signalling, enzymatic function, nutrient uptake, and stress responses, its effects are expected to depend on stimulus intensity. This experiment varied the applied voltage (0 V, 3 V, 6 V, 9 V, 12 V) and stimulated seeds for 10 minutes daily over five days to test how voltage influences early seedling performance. Germination and growth were evaluated using shoot and root lengths as primary outcomes. It was hypothesized that an intermediate treatment (9 V) would produce the greatest positive effect on germination and growth by providing sufficient stimulation without causing injury, whereas 12 V may inhibit growth or progress and lower voltages may be insufficient to elicit measurable improvement[13].

## Methods

Approximately 200 *Zea mays* subsp. *mays* (corn) seeds were first imbibed by soaking in bottled mineral water for 24 hours; floating seeds were removed due to their low likelihood of germination. For each treatment, plastic trays were prepared by lining the base with five layers of thick tissue paper and overlaying a single sheet of filter paper. Each tray was then hydrated with 20 $cm^3$ of water to fully saturate the substrate. Following soaking, 35 seeds were evenly spaced on the moistened filter paper in each tray, and trays were labelled according to the assigned electrical stimulation to avoid cross-contamination between treatments. A 3 V power source was assembled by connecting two 1.5 V batteries in series, and separate 6 V, 9 V, and 12 V batteries were each connected to alligator clip leads. The clips were positioned diagonally opposite one another within each tray, with the positive and negative terminals kept consistently oriented on the same side across all setups. Because electrical exposure parameters are not yet well established for this setup, we used a brief, repeated stimulation schedule of 10 minutes per day over five consecutive days as a practical and low-cost exploratory protocol spanning early germination and initial seedling establishment. To maintain consistent moisture conditions, 10 $cm^3$ of water was added to each tray once per day throughout the stimulation period. After the 5-day treatment, seedling shoot and root lengths were measured using a ruler.

### Data processing

First, germinated seeds were identified from seeds that did not germinate based on the growth of the root and shoot. Presence of either a shoot and root or root only was considered germinated, while the absence of both shoot and root was categorized as ungerminated. To find the percentage germination, the number of germinated seeds was divided by the total seed number and multiplied by one hundred.

$$\text{percentage germination \%} = \frac{\text{number of germinated seeds}}{\text{total number of seeds } (n = 35)} \times 100$$

Next, Microsoft Excel was used to calculate the mean and standard deviation (σ) of the shoot and root lengths of germinated seeds independently.

Then, the germination rate was calculated by dividing the average lengths (i.e. length) by the time given for the seeds to grow in *n* days (n = 5 days).

$$\text{rate of germination} = \frac{\text{length}}{5} = \text{cm/day}$$

### Statistical analysis

Descriptive statistics (mean (μ) and standard deviation (σ)) were calculated in Microsoft Excel, and the error bars in Figure 1 represent ±1 SD. In addition, a chi-square test of independence was used to compare germination outcome (germinated versus ungerminated) across voltage treatments ($\alpha$=0.05).

Finally, to analyze both the root lengths and shoot lengths among the different voltage strengths, a one-way ANOVA test was performed using IBM SPSS Statistics software. The test was utilized to determine whether there were statistically significant differences in the means of root and shoot lengths among the various voltage levels or if it was merely due to random variation. The raw data for root and shoot lengths across different voltage strengths were inputted into the software, and the one-way ANOVA test was conducted assuming equal variance.

Statistical hypotheses

Shoots:

($H_{s0}$): No significant difference in the mean shoot lengths of Zea mays subsp. *mays* seeds is present when subjected to different voltages of electrical stimulation.

($H_{s1}$): A significant difference in the mean shoot lengths of Zea mays subsp. *mays* seeds is present when subjected to different voltages of electrical stimulation.

Roots:

($H_{r0}$): No significant difference in the mean root lengths of Zea mays subsp. *mays* seeds is present when subjected to different voltages of electrical stimulation.

($H_{r1}$): A significant difference in the mean root lengths of Zea mays subsp. *mays* seeds is present when subjected to different voltages of electrical stimulation.

## Results

**Table 1.** Germination outcomes by voltage treatment. Number of germinated and ungerminated seeds, and the resulting germination percentage (%), for seeds exposed to 0 V, 3 V, 6 V, 9 V, and 12 V.

| Voltage (V) | 0 | 3 | 6 | 9 | 12 |
|---|---|---|---|---|---|
| Number of Ungerminated Seeds | 16 | 19 | 16 | 18 | 16 |
| Number of Germinated Seeds | 19 | 16 | 19 | 17 | 19 |
| Percentage Germination (%) | 54.3 | 45.7 | 54.3 | 48.6 | 54.3 |

Shoot and root growth rates across voltage treatments. Mean shoot and root growth rates (cm/day) of germinated corn (*Zea mays* subsp. mays) seeds exposed to electrical stimulation at 0 V, 3 V, 6 V, 9 V, and 12 V. Growth rate was calculated by dividing the mean shoot/root length by the growth period (days). Error bars indicate variability in the underlying length measurements (standard deviation, σ, calculated in Excel).

In Figure 1, mean root growth rate increased from 0 V to 6 V and decreased from 6 V to 12 V. Mean shoot growth rate increased from 3 V to 6 V and decreased from 6 V to 12 V (Figure 1).

The standard deviation of the length values is low for all voltage treatments, with a maximum of 1.34 for the 6 V root sample, indicating that measurements were relatively consistent within each voltage treatment. However, variability was highest for the 6 V roots, thus a less uniform response is present at this voltage.

The root lengths in all of the voltage treatments were greater than those of the shoots. This pattern is consistent with the normal germination sequence, where early root elongation precedes shoot emergence, as reflected by higher growth rates in roots compared to shoots (Table 3). At 6 V, the highest growth rates were observed: the shoot growth rate is 0.24 cm/day, while the root growth rate is 0.4 cm/day (Table 3). However, we can see that the difference between the root and shoot lengths and growth rate is larger around the optimum voltage of 6 V, indicating that electrical stimulation might have a greater effect on the seeds' roots rather than shoots.

Due to the large difference between the roots and shoots results, the ANOVA statistical test was used to determine if the effect of electrical stimulation is significant. Despite observing a difference in the mean lengths

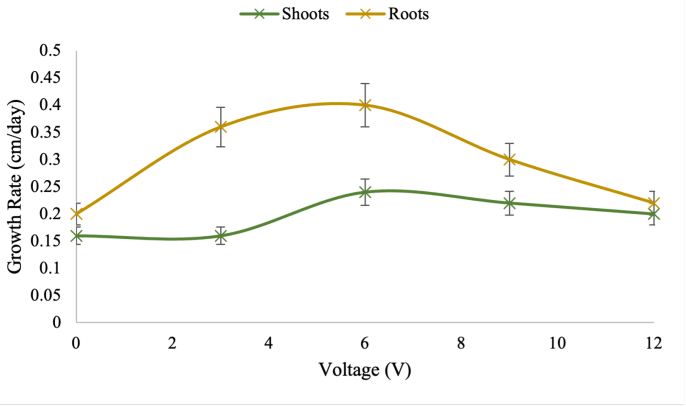



**Figure 1.** Shoot and root growth rates across voltage treatments. Mean shoot and root growth rates (cm/day) of germinated corn (Zea mays subsp. mays) seeds exposed to electrical stimulation at 0 V, 3 V, 6 V, 9 V, and 12 V. Growth rate was calculated by dividing the mean shoot/root length by the growth period (days). Error bars indicate variability in the underlying length measurements (standard deviation, σ, calculated in Excel).

of the shoots at the different voltages, the p-value = 0.410 showed no significance (Table 4). This indicates that $H_{s0}$, suggesting that the observed difference between the voltage strengths for shoots, is not statistically significant. The p-value = 0.014 for the root lengths was significant (Table 4). This suggests that electrical stimulation does not have a significant effect on the growth rate of the shoots, but can positively impact the growth rate of roots at an intermediate voltage of ~6 V.

As observed in Table 1, which details the percentage of germinated seeds (i.e. successful germination), electrical stimulation did not affect percentage germination. This is because the values vary slightly among the voltage levels, with the highest percentage of 54.3% observed at each 0 V, 6 V, and 12 V, and the lowest percentage of 45.7% observed at 3 V (difference of 8.6%). Consistent with this, a chi-square test of independence detected no significant association between voltage treatment and germination outcome ($\chi^2$=0.915, p=0.922, d.f.=4, n=175).

## Discussion

Overall, this experiment indicates that brief electrical stimulation can influence early seedling elongation in corn, with effects that are more apparent for roots than for shoots.

The hypothesis that 9 V would produce the greatest improvement in germination and early growth was not supported. This is shown in Figure 1, where the optimal voltage for both shoot and root growth was 6 V. Compared with 0 V (shoot: 0.80 cm; root: 1.00 cm), the 6 V treatment produced higher mean lengths (shoot: 1.20 cm; root: 2.00 cm), indicating a ~50.0% increase in shoot length and 100.0% increase in root length. The hypothesized optimal condition of 9 V showed smaller increases relative to 0 V (shoot: 1.10 cm, +37.5%; root: 1.5 cm, +50.0%). Across treatments, growth followed a non-linear pattern: elongation increased from 0-6 V and declined at higher voltages of 9 and 12 V, consistent with an "optimal stimulation window" rather than a strictly linear dose-response. Overall, the trend resembles a downward-opening parabola, and it suggests a relatively symmetric pattern around the vertex at 6 V (more pronounced in roots than in shoots). However, given the variability within groups (Table 2), this pattern should be interpreted cautiously and confirmed with additional replication and longer growth periods. Standard deviations were generally low, suggesting the pattern is not driven by a small number of extreme values.

Analysis of the growth rates reveals that electrical stimulation has a more

**Table 2.** Mean shoot and root lengths of germinated seeds by voltage treatment. Mean shoot and root lengths (cm) for germinated seeds at 0 V, 3 V, 6 V, 9 V, and 12 V, with corresponding standard deviations (σ). Mean lengths are reported with a measurement precision of ±0.1 cm.

| Voltage (V) | 0 | | 3 | | 6 | | 9 | | 12 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Shoot | Root | Shoot | Root | Shoot | Root | Shoot | Root | Shoot | Root |
| Mean Length (±0.1cm) | 0.8 | 1.0 | 0.8 | 1.8 | 1.2 | 2.0 | 1.1 | 1.5 | 1.0 | 1.1 |
| Standard Deviation ($\sigma$) | 0.81 | 0.85 | 0.86 | 0.84 | 0.97 | 1.34 | 0.72 | 0.80 | 0.68 | 0.85 |

**Table 3.** Shoot and root growth rates by voltage treatment. Shoot and root growth rates (cm/day) for germinated seeds at 0 V, 3 V, 6 V, 9 V, and 12 V. Growth rate was calculated by dividing the mean shoot/root length by the growth period (days).

| Voltage (V) | 0 | | 3 | | 6 | | 9 | | 12 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Shoot | Root | Shoot | Root | Shoot | Root | Shoot | Root | Shoot | Root |
| Growth rate (cm/day) | 0.16 | 0.20 | 0.16 | 0.36 | 0.24 | 0.40 | 0.22 | 0.30 | 0.20 | 0.22 |

prominent effect on root growth compared to shoots. Statistical analysis also shows that differences in root lengths are significant, whereas those of the shoots are not. Specifically, one-way ANOVA indicated a significant effect of voltage on root length ($p$=0.014) but not on shoot length ($p$=0.410) (Table 4), implying that electrical stimulation selectively impacted early root elongation under the conditions tested. The root–shoot difference was largest around 6 V, suggesting stimulation may have a stronger effect on roots than shoots near the apparent optimum. The percentage of germinated seeds is relatively consistent across different voltage levels (45.7-54.3%), suggesting that electrical stimulation did not have a meaningful change in germination success in this experimental setup.

This study addressed how varying the applied voltage (0-12 V) affected germination and early growth, measured by shoot and root length. The results support an effect on root growth: root length differed across voltage treatments (one-way ANOVA, p = 0.014), with the highest mean observed at 6 V (2.00 cm). The effect on germination success rate and shoots was, however, not statistically significant in this dataset.

Insights into the effect of electrical stimulation on corn seeds have potential implications in the field of agriculture. The observed improvement in root growth under specific electrical voltages suggests an innovative approach to improving crop development, particularly in the early stages that are crucial for yield. Notably, this outcome is partially consistent with prior work showing that electrical stimulation can enhance maize seedling growth, while higher-intensity exposure may become inhibitory depending on stimulation parameters and stress responses[9]. However, further research is essential to understand the long-term impacts on mature plants and overall yield, especially for vital crops like corn. In natural or *in situ* settings, crops may also experience electrical or mechanical stimuli (e.g., ambient electric fields near infrastructure, wind, vibration, etc.) and determining whether these exposures produce comparable effects remains an important direction for future work.

## Limitations

The effective sample size per treatment was limited by substantial non-germination. Although 35 seeds per voltage were used, approximately half did not germinate, reducing statistical power and the robustness of between-group comparisons; additionally, the five-day duration may not have provided sufficient time for all seeds to complete the germination process. Only one experimental run was completed, eliminating statistical inferences over replicates. The seeds used were not of the same size, meaning nutrient reserves likely differed between seeds, which can affect the germination rate and process. In addition, the use of tissue and filter paper as a substrate did not provide the seed with the essential nutrients to grow. The measuring instrument had limited precision (reported as ±0.01 cm), which might not have correctly identified more subtle differences. Additionally, the shoot and root of the corn seeds were usually bent, which made measuring their lengths difficult (consequently, measurements were taken from the tip to the bend and from the bend to the end). Roots were sometimes branched; thus, the longest branch was measured. Together, these issues limit construct validity because one-dimensional length measurements do not capture traits such as branching architecture, total root surface area, or biomass, which may respond differently to stimulation. In terms of statistics, only an omnibus one-way ANOVA was performed, and post-hoc tests (e.g., Tukey HSD) would be needed to determine which voltage treatments differ from each other.

Even though seeds were acquired from the same source, confounding variables such as seed size can still exist and affect the results. To prevent this, seed sizes can be controlled using either a balance or a ruler. Providing the corn seeds with nutrients (soil or ½-strength Murashige and Skoog nutrient medium) would improve ecological relevance and may reduce nutrient-limitation effects on growth[14]. A climate chamber can be used to control factors like temperature, moisture and air quality. Regular checks and recalibrations of the batteries would be crucial for maintaining the accuracy of voltage delivery. More importantly, recording the delivered current (mA) and electrode spacing would better standardize exposure across trays.

**Table 4.** One-way ANOVA results for shoot and root lengths across voltage treatments. One-way ANOVA summary statistics for shoot and root length measurements across five voltage conditions at 0 V, 3 V, 6 V, 9 V, and 12 V, including reporting sum of squares, degrees of freedom (df), mean square, F-statistic (F), and significance (Sig.). For shoots, the ANOVA returned F=1.004 with p-value =0.410; for roots, F=3.328 with p-value =0.014.

| Shoots | | | | | |
|---|---|---|---|---|---|
| | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 2.663 | 4 | .666 | 1.004 | .410 |
| Within Groups | 56.353 | 85 | .663 | - | - |
| Total | 59.016 | 89 | - | - | - |
| | | Roots | | | |
| | Sum of Squares | df | Mean Square | F | Sig. |
| Between Groups | 12.571 | 4 | 3.143 | 3.328 | .014 |
| Within Groups | 80.274 | 85 | .944 | - | - |
| Total | 92.845 | 89 | - | - | - |

The effect of varying electric voltages can be evaluated more accurately by measuring the dry mass of the germinated seeds. This would overcome the issue of root branching. Ensuring consistent seed spacing is essential and verifying that the electrical field experienced by seeds is comparable within and across trays (rather than relying only on nominal battery voltage) would strengthen internal validity.

## Future directions

Future directions include investigating the long-term effects of electrical stimulation on mature corn plants, because this experiment does not elucidate the effect of electrical stimulation on the crop yield of the maize plant. A longer study should track plants through vegetative growth and reproduction while testing whether early root elongation at 6 V translates into measurable agronomic outcomes. Since maize is a major food crop, it is essential to determine whether electrical stimulation improves productivity under realistic conditions. Therefore, assessing outcomes such as cob size, kernels per cob, and total biomass (with appropriate replication and field- or greenhouse-relevant substrates) would clarify whether electrical stimulation produces durable benefits or only short-term changes in early development.

## Acknowledgements

The author would like to sincerely thank Lara Nahas Feghali and Juhaina Abou Fadel for their valuable support throughout the development of this study. Their combined expertise in biology and laboratory practice provided important guidance and encouragement during both the experimental and writing phases of this research.

## References

1. Reed, R., Bradford, K. & Khanday, I. Seed germination and vigor: ensuring crop sustainability in a changing climate. *Heredity* **128,** 450–459 (2022). 10.1038/s41437-022-00497-2
2. Grzelka, K., Matkowski, A. & Ślusarczyk, S. Electrostimulation improves plant growth and modulates the flavonoid profile in aeroponic culture of Scutellaria baicalensis Georgi. *Front. Plant Sci* **14,** 1142624 (2023). 10.3389/fpls.2023.1142624
3. Footitt, S., Huang, Z., Ölcer-Footitt, H., Clay, H. & Finch-Savage, W. The impact of global warming on germination and seedling emergence in Alliaria petiolata, a woodland species with dormancy loss dependent on low temperature. *Plant Biol. (Stuttg* **20,** 682–690 (2018). 10.1111/plb.12720
4. Alhassan, G., Aliyu, J., Jokthan, G. & Musa, A. *Plant Growth and Development* (NOUN Press, 2019).
5. Carrera-Castaño, G., Calleja-Cabrera, J., Pernas, M., Gómez, L. & Oñate-Sánchez, L. An updated overview on the regulation of seed germination. *Plants (Basel* **9,** 703 (2020). 10.3390/plants9060703
6. Finch-Savage, W. & Leubner-Metzger, G. Seed dormancy and the control of germination. *New Phytol* **171,** 501–523 (2006). 10.1111/j.1469-8137.2006.01787.x
7. Blacklow, W. Influence of temperature on germination and elongation of the radicle and shoot of corn (Zea mays L. *Crop Sci* **12,** 647–650 (1972). 10.2135/cropsci1972.0011183X001200050028x
8. Fromm, J. & Lautner, S. Electrical signals and their physiological significance in plants. *Plant Cell Environ* **30,** 249–257 (2007). 10.1111/j.1365-3040.2006.01614.x
9. Li, Z.-G., Gou, H.-Q. & Li, R.-Q. Electrical stimulation boosts seed germination, seedling growth, and thermotolerance improvement in maize (Zea mays L. *Plant Signal. Behav* **14,** 1681101 (2019). 10.1080/15592324.2019.1681101
10. Hodges, D., DeLong, J., Forney, C. & Prange, R. Improving the thiobarbituric acid-reactive-substances assay for estimating lipid peroxidation in plant tissues containing anthocyanin and other interfering compounds. *Planta* **207,** 604–611 (1999). 10.1007/s004250050524
11. Gill, S. & Tuteja, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem* **48,** 909–930 (2010). 10.1016/j.plaphy.2010.08.016
12. Bailly, C. The signalling role of ROS in the regulation of seed germination and dormancy. *Biochem. J* **476,** 3019–3032 (2019). 10.1042/BCJ20190159
13. Mamlic, Z. The use of electrostatic field to improve soybean seed germination in organic production. *Agronomy* **11,** 1473 (2021). 10.3390/agronomy11081473
14. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant* **15,** 473–497 (1962). 10.1111/j.1399-3054.1962.tb08052.x

Submitted: 01/15/2026
Accepted: 02/20/2026
Published: 04/10/2026

Research Article

[1] Department of Biology, McGill University, Montreal, QC, Canada
[2] Rubenstein School of the Environmental Sciences, University of Vermont, Burlington, VT, United States
[3] Agricultural and Environmental Sciences, McGill University, Montreal, QC, Canada

Keywords

crayfish, morphometric analysis, species identification, habitat effects, freshwater ecosystems

Email Correspondence

corinne.lapierre@mail.mcgill.ca

https://doi.org/10.26443/msurj.v21i1.406



Corinne Lapierre[1], Marlee Bickerdike[1], Kate Healy[2], Karim Chehata[3]

# Species Composition and Morphological Variation of Crayfish in the Gault Nature Reserve

## Abstract

Crayfish play a crucial role as ecosystem engineers in freshwater ecosystems, yet they remain poorly represented in ecological studies. Integrating aquatic invertebrates into biodiversity monitoring efforts is urgently needed, particularly given their important ecological role as indicators of freshwater ecosystem health. The purpose of this pilot study was to record crayfish presence in the Gault Nature Reserve and to characterize species identity, morphology, and dorsal colouration of individuals from two habitats sampled (Lake Hertel and a nearby stream), with colouration assessed for substrate matching. This study records the occurrence of two crayfish species in the Gault Nature Reserve of McGill University and examines their morphological and colouration differences. Two species of crayfish were identified: the Virile crayfish (*Faxonius virilis*) in Lake Hertel and the Appalachian Brook crayfish (*Cambarus bartonii*) in a nearby stream. A key outcome of this pilot survey is the identification of different crayfish species living in different habitat types within the reserve. This finding demonstrates the necessity of broader biodiversity surveys of freshwater species in Gault. Morphometric analyses showed significant size differences between species aligning with ecological differences observed between the two site-specific species, though these disparities likely reflect species–specific traits more than habitat effects alone, which could not be isolated in this study design. Colouration analyses did not support the hypothesis that crayfish colouration matches substrate colours for camouflage. The analysis of colouration is considered exploratory and must not be considered a conclusive test of substrate matching because of the failure of photographic colour extraction to match colouration and the absence of consideration of body size, sex, and developmental stages. Future research should incorporate more controlled colour assessments, expand the survey area, and conduct genetic analyses and transplant experiments to clarify whether colouration and morphology are determined by environmental or genetic factors.

## Introduction

Fully documenting all components of an ecosystem is often limited by the scale of resources needed. However, failing to sufficiently study all the components of an ecosystem can result in blind spots that compromise restoration efforts, species protection[1], and invasive species management. The Gault Nature Reserve of McGill University located in Mont St–Hilaire, Quebec, Canada, is a protected reserve of 1,000-hectare of old-growth forest within Canada's first UNESCO Biosphere Reserve, where a sugar maple-hickory forest matrix provides a variety of terrestrial and freshwater habitats. Although a significant body of work has been conducted, freshwater ecosystems are understudied relative to terrestrial habitats at the reserve, a trend observed worldwide [2]. Freshwater biodiversity is declining much faster than most terrestrial ecosystems, yet it remains significantly underrepresented in research and conservation efforts[1, 3]. As the quantity of research projects and conservation efforts increases at the Gault Reserve, it is essential to broaden the research focus to develop a comprehensive understanding of the mountain's ecology. Additionally, when engaging in invasive species management, recreational activities, scientific research, and trail maintenance that can affect aquatic life, there is a critical need to have a full account of the organisms living on the mountain[4, 5]. Furthermore, as the reserve increases its efforts in invasive species management, currently with terrestrial plants, it is crucial to ensure invasive species, like the Rusty crayfish (*Faxonius rusticus*), which can cause major disruptions in freshwater ecosystems[6], are detected early in aquatic habitats. Historical gaps in surveillance mean that no mollusks or crustaceans have been officially recorded by the reserve, although past anecdotal reports suggest the presence of both. In contrast, over 200 species of birds and more than 800 species of microlepidoptera have been officially recorded.

Crayfish are critical components of freshwater systems because they are omnivorous, serve as prey to many fishes, and function as ecosystem engineers[7, 8]. Crayfish alter their habitats by changing detrital processing rates, causing bioturbation of sediments, and altering algal cover on substrates[7, 9]. Moreover, crayfish are considered keystone controllers of trophic webs, bioindicators of ecosystem health, and indicator species of environmental conditions such as pollution and/or water temperature changes[8]. In headwater streams, crayfish can increase the rate of leaf litter decomposition, as well as affect the amounts and distributions of fine particulate matter, thus affecting benthic habitat conditions[9]. Crayfish can also cause bioturbation of stream beds, thus inhibiting particle consolidation, affecting algal cover, as well as sand-gravel surface composition[7].

Crayfish traits vary across habitats, and both morphology and colouration can reflect interacting biological and environmental factors[8, 10]. Body size and shape can covary with flow conditions and substrate structure[10, 11]. Research has demonstrated habitat-size and shape variation in crayfish, including Perry et al. (2013), which showed variation in rusty crayfish morphology based on lake and stream habitat with differing flow rates, and Clark et al. (2008), which showed size-dependent habitat use among lotic crayfish, with larger individuals tending to occur in deeper, slower pools and smaller individuals tending to be found in riffles with coarser substrate. Colour phenotypes may correlate with habitat background, and

colouration can also be influenced by diet and the breeding season; however the functional significance of colouration in freshwater crayfish remains unclear[12, 13].

This study aims to address this knowledge gap by providing the first evaluation of crayfish presence, identification, and morphology in the Gault Nature Reserve. Specifically, we aim to: [1] confirm the presence of crayfish in Lake Hertel and a nearby stream and, if present, identify them to species and [2] compare their morphology and colouration differences in relation to their distinct habitats. We hypothesized that crayfish occur in the reserve and that individuals from different habitats will differ in morphology. We expect that morphological characteristics such as carapace, abdomen, and claw length or width, as well as total length, will be smaller in crayfish from stream habitats compared to lake habitats due to potential adaptations to distinct environmental conditions related to habitat types, such as substrate type and colour, water flow, depth, and predator and prey types. In particular, crayfish from faster-flowing habitats may exhibit smaller morphologies than crayfish from slower or lentic habitats. Additionally, we expect that both groups' dorsal exoskeleton colouration will match their respective habitat substrate colour to improve camouflage ability.

Since each species is represented at only one sampling site, habitat type and species are confounded in this pilot study. We therefore regard all morphology and colour data as descriptive data associated with each site rather than as tests for habitat effect. Our main goal is to record the occurrence and species identity of crayfish at two freshwater habitats in the reserve, as well as the associated morphometrics and dorsal colour patterns for the individuals collected.

## Methodology

As, to our knowledge, the presence of crayfish in the Gault Nature Reserve of McGill University has never been officially documented, we relied on anecdotal reports to identify potential site locations. A pilot study was done on July 10, 2025, to determine our capacity to find and catch crayfish with different capture techniques in 3 locations (2 streams and Hertel lake). For this study, 2 of the 3 sites identified in the pilot study were retained for this study and were sampled for crayfish on July 11 and 12, 2025, in the Gault Nature Reserve of McGill University, located in Mont St–Hilaire, Quebec, Canada.

We captured crayfish by visually locating the specimens and catching them manually. Each site was sampled three times over a two–day period. A total of 41 crayfish were retained for experimentation. The sample size from the lake site was 28 crayfish (*Faxonius virilis*) (18 males, 9 females, 1 individual of unidentifiable sex) and 13 from the stream site (*Cambarus bartonii*) (1 male, 9 females, 3 individuals of unidentifiable sex). Life stages were not recorded. Crayfish were sedated with clove oil and ethanol then transferred to a shallow beige container to provide a uniform background. Photographic lights and a tripod holding an iPhone were used to take dorsal and ventral photographs. Crayfish were marked with nail polish before release to avoid recapture.

Photographs and observations were taken using a WILD M3C microscope (Heerbrugg, Switzerland) to locate unique features that can be used for species identification, including the presence of hairs and/or spines on specific body parts, rostrum shape, etc. The crayfish we found in the lake were identified as *Faxonius virilis* (The Virile crayfish). We identified the stream crayfish as *Cambarus bartonii* (Appalachian Brook crayfish). Crayfish morphometrics (carapace length [CL], carapace width [CW], abdomen length [AL], and total length [TTL]) were analyzed in R, testing site/species–associated differences with MANOVA, then ANOVA. Photos were processed to extract HEX colours, converted to CIE Lab, and compared with substrate reference colours using ΔE (CIEDE2000); since the extracted colours were generally darker than the photographed appearance, the colour analysis was considered preliminary.

A more thorough description of the methodology is presented in the supplementary materials (S1). The dataset (S2) and R scripts (S3–5) are available as supplementary materials.

## Results

Two species of crayfish were found and identified at the Gault Nature Reserve of McGill University. We identified the stream crayfish as the Appalachian Brook crayfish (*Cambarus bartonii*) and the Lake Hertel crayfish as the Virile crayfish (*Faxonius virilis).*

Lake crayfish (*Faxonius virilis*) were significantly larger than stream crayfish (*Cambarus bartonii*) across all four body measurements combined (MANOVA: Pillai=0.649, $F$=17.59, df=4, 38, $p<3.02\times10^{-8}$). Univariate ANOVAs confirmed Lake crayfish (*Faxonius virilis*) were significantly larger than stream crayfish (*Cambarus bartonii*) for carapace length (ANOVA: $F$=32.13, df=1, 41, $p<1.28\times10^{-6}$), carapace width ($F$=17.81, df=1, 41, $p<1.32\times10^{-4}$), abdomen length ($F$=19.43, df=1, 41, $p<7.36\times10^{-5}$), and total length ($F$=32.26, df=1, 41, $p<1.24\times10^{-6}$) (Figure 1A). Because species and site are confounded, this difference reflects species/site differences rather than a tested habitat effect. Univariate ANOVAs were interpreted conditional on a significant MANOVA; p–values were additionally adjusted using the Holm procedure to correct for multiple comparisons. All univariate effects remained significant when Holm correction was applied. Effect sizes and 95% CI for all traits and Holm–adjusted p–values are presented in Supplementary Material S1. All data met the assumptions required for MANOVA and ANOVA

Carapace length was 24.00 ± 0.73 mm in the lake group and 16.24 ± 1.27 mm in the stream group; carapace width was 10.52 ± 0.32 mm and 7.81 ± 0.64 mm; abdomen length was 15.74 ± 0.53 mm and 10.96 ± 1.12 mm; and total length was 43.35 ± 1.16 mm and 30.41 ± 2.24 mm, respectively.

Although no statistical test was conducted, box plots suggest that males tended to be larger than females across all four body measurements (Figure 1B). We note that the lake crayfish group included mostly males, while the stream group was mostly female, which may influence size comparisons. Because life stage and sex were not included as covariates, the observed patterns should be interpreted cautiously.

Lake crayfish (*Faxonius virilis*) had a significantly smaller colour difference (ΔE) from the other habitat substrate compared to their own (ANOVA: $F$=5.258, df=1, 56, $p$=0.0256) (Figure 2). Based on extracted colours, Lake crayfish (*Faxonius virilis*) had a dorsal exoskeleton colouration more similar to the stream substrate than to the lake substrate. Stream crayfish (*Cambarus bartonii*) showed no significant difference in colour similarity between their own and the other habitat substrate ($F$=0.3061, df=1, 24, $p$=0.585) (Figure 2). However, since the extracted HEX values failed to correspond with the photographed appearance, the colour results are seen as exploratory and are not taken as a definitive test of substrate matching.

## Discussion

In this study, we sampled one location in a lake habitat, Lake Hertel, and one stream location to confirm crayfish presence, identify the species, and compare their morphology and colouration differences in relation to their distinct habitats. We hypothesized that if crayfish are present in the Gault Nature Reserve, individuals from different habitats would differ, and all mor-

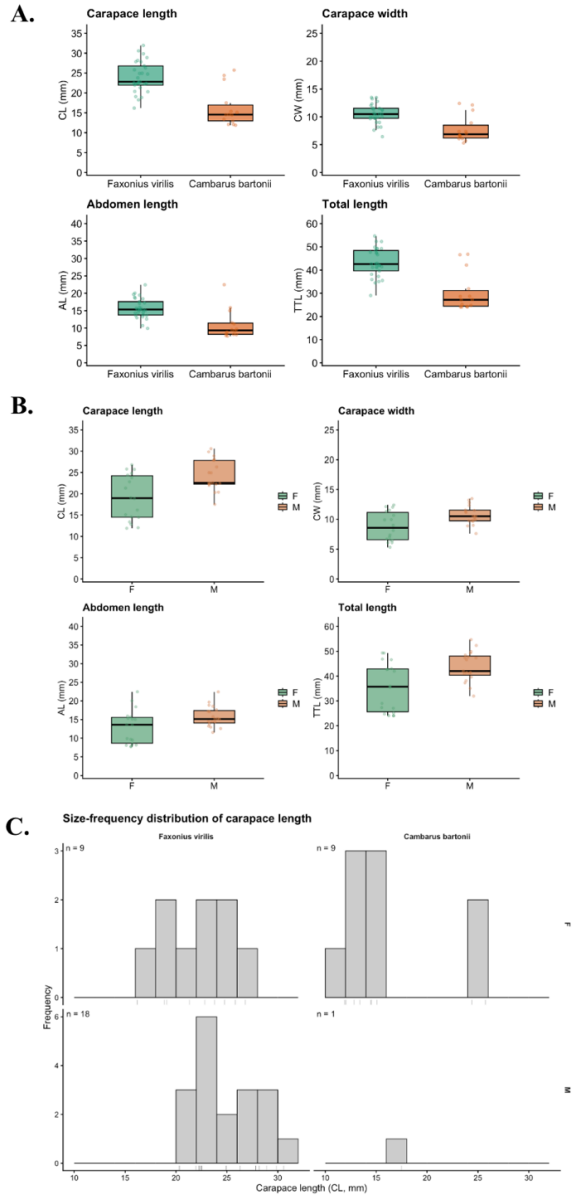



**Figure 1.** Boxplots comparing body measurements (carapace length, width, abdomen length, and total length) (A) Between Lake Hertel and stream crayfish. Lake specimens (Faxonius virilis) were significantly larger across all metrics than stream crayfish (Cambarus bartonii) (MANOVA: Pillai = 0.649, F = 17.59, df = 4.38, $p < 3.02 \times 10^{-8}$). (B) Body measurements of Lake Hertel and stream crayfish grouped by sex. Boxplots show carapace length, width, abdomen length, and total length, illustrating size variation by sex and between the two site-specific species. Because habitate and species are confounded, differences cannot be attributed to habitat alone. (C) Size-frequency histograms of carapace length (CL, mm) shown by species and sex, with sample sizes indicated in each panel, provided to visualize the observed size distributions and potential size class structure. Histograms are descriptive only because sex ratios were unbalanced and life stage was not recorded.

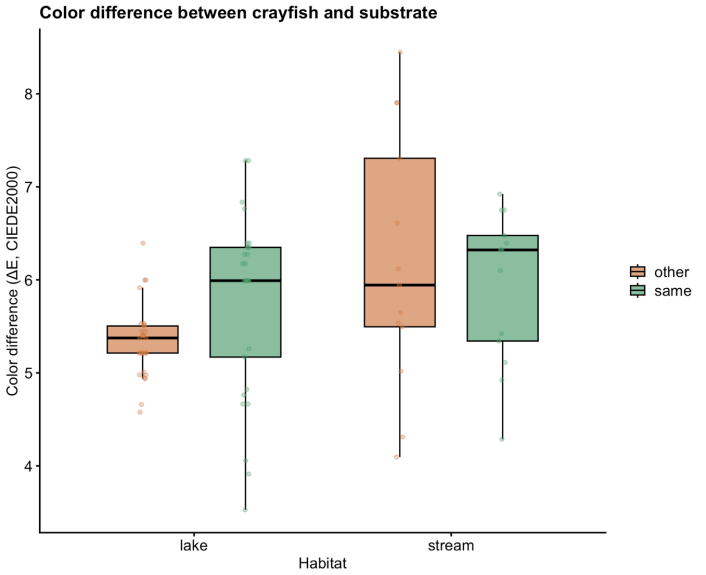



**Figure 2.** Color difference (ΔE, CIEDE2000) between crayfish dorsal exoskeleton and substrate from both habitats. Lake crayfish (Faxonius virilis) showed a significantly smaller ΔE when compared to the stream substrate than to their own (ANOVA: F = 5.258, df = 1, 56, p = 0.0256). Stream crayfish (Cambarus bartonii) showed no significant difference between substrates. However, this finding is considered preliminary because of the discrepancy between the hex values obtained and what was photographed.

phology metrics of stream crayfish (*Cambarus bartonii*) would be smaller than those of lake crayfish (*Faxonius virilis*) and that both groups' dorsal exoskeleton colouration would more closely match their respective habitat substrates.

Insufficient inventories of freshwater species create critical blind spots for conservation, invasive species management, and early detecting of ecological change [1, 3]. Species identification revealed not one but two different species in the two sites sampled at the Gault Nature Reserve: the Appalachian Brook crayfish in the stream habitat and the Virile crayfish in the lake habitat. When this study was initiated, variations in morphology were expected to be driven by habitat differences as we only expected to find one species of crayfish based on anecdotal reports. Species identification fundamentally changes the interpretation of the morphometric analyses. Since each species of crayfish was represented at only one site, species identity and habitat are completely confounded in this experiment. Therefore, habitat effects cannot be separated from baseline interspecific differences without replication of habitats within species. These results can be considered a good baseline with regards to detection records from two sampled locations. However, it must be noted that the results must be considered as presence records rather than absence records for both species in other habitats within the reserve.

The most significant finding of this investigation is not the preliminary colour result, but the discovery of two crayfish species in different habitat types in the Gault Reserve. This indicates habitat filtering, habitat specificity of the two crayfish species, or biotic interactions (such as competition and predator-mediated habitat use) as possible explanations for the observed segregation of the two crayfish species. Since each of the two crayfish species was collected in only one habitat type in this preliminary investigation, future investigations in the reserve and seasonally replicated investigations must verify whether this segregation is consistent through time and whether the two crayfish species co-occur in other areas of the reserve.

Adult Appalachian Brook crayfish can range from 18.5 mm to 36.7 mm in carapace length, while the Virile crayfish can range from 23.2 to 55.0 mm [14, 15]. In our study, the Appalachian Brook crayfish (maximum cara-

pace length was 25.74 mm) from the stream were smaller by all metrics compared to the Virile crayfish (maximum carapace length 29.86 mm), although our data was not separated by developmental stages or sex (Figure S3B). Since juveniles were not identified and excluded, size and color estimates may reflect a mixture of life stages, which may cause bias within and between sites. Although size difference likely reflects species–related differences, results still align partially with our hypothesis that stream crayfish (*Cambarus bartonii*) would be smaller than lake crayfish (*Faxonius virilis*), consistent with ecological patterns observed in other systems where body size correlates with habitat constraints such as water depth, flow, and resource availability. Perry et al. (2013) demonstrated this intraspecific trend in Rusty crayfish (*Faxonius rusticus*), which were significantly smaller in high–velocity streams compared to low–velocity streams and lakes[11]. Furthermore, Clark et al. (2008) found that the Allegheny crayfish (*Faxonius obscurus*) of larger size tend to occupy deeper pools with slower current velocity and smaller grain sizes, while smaller individuals preferred riffle habitats with substrates made of coarser grain size[10]. For a direct comparison, reciprocal transplant experiments can be conducted in the future by rearing juvenile stream crayfish (*Cambarus bartonii*), under lake-like conditions and juvenile lake crayfish (*Faxonius virilis*), under stream-like conditions to assess whether the size differences persist under different environmental conditions. This type of experiment would be possible as, although both species were found at the Gault Nature Reserve only in their preferred environment, both species are known to live in both lake and stream environments[14]. Body size is ecologically significant in crayfish because it affects resource usage, food and shelter competition, species interaction, and susceptibility to variable predation[8]. The pronounced differences observed between species at Gault suggest strong species-habitat segregation (or habitat association) within the reserve, which could reduce interspecific crayfish competition and stabilize the community structure. Recognizing these species–specific patterns is essential for conservation, in order to ensure that conservation efforts are not compromised by the failure to identify or monitor all species present, which could result in ineffective or potentially harmful management strategies.

The survival and evolutionary success of many organisms largely depend on their capacity to avoid detection by predators[16]. The second part of our hypothesis predicted that crayfish dorsal colouration would closely match the substrate of their native habitat, reducing predation risk, an ability many crustaceans have evolved[16, 17]. To our surprise, the statistical analysis did not support this prediction: neither species showed a statistically significant match to their own substrate compared to the substrate from the alternate habitat. In fact, the Virile crayfish matched the colour of the stream substrate better than its own substrate according to our data. This result did not support the camouflage hypothesis, and the colour findings are considered preliminary due to the limitations in the image–based colour extraction. Given that the substrate reference colours were based on a habitat average and that the crayfish colours were based on an extraction from all non–background pixels without glare masking, the colour analysis is interpreted as a preliminary analysis rather than an actual test of camouflage. Although multiple mechanisms could contribute to colouration, the current data do not allow strong inference about selection or signaling, so the colour component is interpreted as exploratory. It is possible that background matching is not a dominant selective pressure for these species in the habitats we studied, as both the studied species used shelter, mostly rocks, to hide during capture, which may indicate a greater reliance on physical refuge than on colouration for predator avoidance. However, given the preliminary nature of the obtained results, interpretations remain speculative. Freshwater crayfish show variations in colour within species, populations, and even intra–individual differences, but the function behind dorsal colouration is poorly understood[18]. The colour discrepancy between the substrate and the dorsal exoskeleton could be explained by colour variation that occurs throughout their life stage. Indeed, colour can vary depending on developmental and reproductive stage in some crayfish species [15, 17, 19]. Additionally, some studies suggest that crayfish colouration can also be influenced by diet [19, 20]. Our results underscore that colouration in crayfish may result from multiple interacting factors, including ontogeny, diet, and reproductive strategies, rather than camouflage alone, if it has any function at all although, our results being preliminary, may not accurately capture true colouration patterns

However, it is possible that the results from this report may not accurately represent true crayfish colouration of the studied population. The obtained HEX colour values represent mostly shades of black when visualized which contradicts our observations, where the crayfish collected were lighter and showed more inter–individual colour variation (Figure S2D–E, Figure S2I–J, Figure S3A–C). This was especially true for the Virile crayfish that produced very dark values while in the laboratory the Appalachian Brook crayfish adults were the darkest (Figure S2D–E, Figure S2I–J, Figure S3C–IDS6). All team members noted that captured crayfish appeared to better match the substrate from their own habitat, with the stream's rocky substrate and associated crayfish having more reddish tones, while the lake crayfish had more neutral colours and spots that blended better into the mix of sand, pebbles, and rocks of the lake (Figure S4). The lake crayfish (*Faxonius virilis*) had more neutral colours and spots that blended better in the mix of sand, pebbles, and rocks of the lake (Figure S4). These colour values may be due to suboptimal lighting and image standardization that compressed the colour values and tended to favor darker colours. It is possible that the Virile crayfish matched the colour of the stream substrate better than its own lake substrate is due to overlapping colour ranges between the two habitats. It may also be a function of the crayfish sample size being too small, rather than representing an adaptive response to a substrate these crayfish do not inhabit. While methodological issues such as those described above cannot be excluded, there is some support in the literature for our initial substrate colour matching hypothesis. Nevertheless, this support from the literature does not address the discrepancy in our own results. While Mathews (2025) demonstrated that ontogenetic changes in the Virile crayfish are strongly correlated with colour, this pattern of strong inter– and intra–individual colour variation occurred in the fall, not the summer[17]. According to Mathews (2025), during the fall reproductive season, inter–individual colour variation is high, with some crayfish developing bright green or blue chelae while others remain brown, and intra–individual variation is also high as chelae differed in colour from the carapace and abdomen, both of which did not correlate with sex or body size[17]. In contrast, during the summer when our study took place, most individuals displayed uniformly brown colouration (Mathews, 2025) [17], meaning these seasonal reproductive color changes were less likely to be the primary source of variation in our summer samples. In light of the above, we believe that the most likely explanation for the discrepancy between observed color and extracted color values in this dataset is methodological limitations in image standardization and photographic color extraction, although biological color variation cannot be ruled out.

Despite these methodological limitations, previous studies on Virile crayfish provide valuable context that can help interpret our findings. Using laboratory and field experiments, Thacker et al. (1993) found that the colour of Virile crayfish strongly correlated with its habitat substrate colour, but diet did not, which is consistent with our hypothesis. Additionally, field transfer experiments demonstrated the ability of adults to change colour[12]. Light penetration measurements suggest that light wavelengths available in a habitat are related to the types of colour morphs present[12]. Moreover, Thacker et al. (1993) found that crayfish color changed in a laboratory rearing experiment in which juvenile *Orconectes virilis* were held in separate lake waters, indicating that habitat-related water conditions may play a role in crayfish color expression[12]. However, various experiments showed no difference in survival rates of colour morphs regardless of whether they matched their substrate or not, suggesting that it is still unclear whether colour serves any function to the Virile crayfish. Seasonal and body region

colour variation has been reported for *Faxonius virilis*, but the data do not allow inference regarding signaling versus camouflage because of the limitations of photographic colour extraction.

The literature review for Appalachian Brook Crayfish yielded limited and sometimes inconsistent information on coloration. Fewer studies were found that directly address dorsal exoskeleton colour for this crayfish species, potentially reflecting the confusion and disagreement in the literature regarding this species' taxonomic ID. Some experts have proposed subspecies (*C. b. bartonii, C. b. cavatus, C. b. carinirostris*), but this level of taxonomic resolution is beyond the scope of this paper [21, 22]. The broad spread in ΔE values from this study suggests that the Appalachian Brook Crayfish in the stream had a wider range of colours than the Virile Crayfish from Hertel Lake. As we did not separate crayfish by developmental stage for our colour analysis, ontogenetic differences in colour may explain this broad spread, but to date no correlation between ontogenetic changes and colour has been found elsewhere in the Appalachian Brook crayfish, indicating that intraspecific colour variation is likely due to other factors[23]. However, the Massachusetts Division of Fisheries and Wildlife (2025) states that younger individuals of Appalachian Brook crayfish are lighter, which is also what we observed in the field (Figure S2I–J)[24]. No study was found addressing whether this species tends to colour match its substrate, which points to a gap in the literature.

The Bartlett's test revealed a highly significant difference in variance for colour difference values ($p \approx 1.02 \times 10^{-7}$), due mostly to the Appalachian Brook stream crayfish having a much higher colour variance (Figure 2). This broad spread in ΔE values suggests that the Appalachian Brook crayfish had a wider range of colours than the Virile crayfish from Hertel Lake. Several explanations could explain this pattern. One possibility would be ontogenetic variation, as we did not separate crayfish by developmental stage for our colour analysis, but our review revealed disagreement among sources regarding the correlation between ontogeny and Appalachian Brook crayfish colouration[23, 24].

Another explanation is hybridization, as the Appalachian Brook crayfish is known to form hybrid zones with related species, potentially introducing colour polymorphisms[25], explaining the wider colour range reflected in our results. Although less likely, as we only detected one species of crayfish in this stream (*Cambarus bartonii*), larger surveys and DNA analyses would be needed to exclude this theory. Alternatively, this variance in dorsal colour may reflect deeper evolutionary forces. Graham (2023) proposed that crayfish colours may be a neutral trait caused by genetic drift or pleiotropic effects rather than adaptations for camouflage[13]. Under this hypothesis, high colour variability could occur because it offers neither advantage nor disadvantage in terms of fitness. This would explain the high variance from our results and why our data showed no clear substrate–matching pattern despite observed camouflage during fieldwork. Ultimately, the pronounced colour variability observed in stream crayfish (*Cambarus bartonii*) may result from factors such as hybridization or evolutionary processes unrelated to strict substrate matching, though adaptive background–matching should not yet be ruled out.

Future studies should aim to address limitations of this study by separating individuals by sex and developmental stage for morphometric and colour analyses, as variation might be at least partly due to ontogeny. In addition, replicating this study with a higher quality camera, superior lighting, stronger sedative to keep crayfish still, standardized conditions, with consistent diet and at different reproductive seasons would help clarify the biological significance of colouration in the Virile crayfish and the Appalachian Brook crayfish. Transplant experiments between lake and stream habitats could reveal whether observed body size differences are environmentally induced or species–specific and whether dorsal colour would change to match the substrate. Expanding surveys across all water bodies in the reserve is needed, as they were not covered by this study, combined with genetic testing (e.g., DNA barcoding) to confirm species identification and rule out potential hybridization. This is especially important for conservation, with invasive crayfish species being a threat in Quebec and some even being able to hybridize with native species in some cases [25, 26, 27]. Future studies should also test what accounts for the segregation of these two species into the lake and stream, respectively, by investigating if they differ in habitat preference, seasonal movement, or competitive interactions that restrict co-occurrence. Ultimately, understanding species identity, morphology, and colouration has real implications for conservation planning. Crayfish are important prey, omnivores, and ecosystem engineers; changes in their traits or abundance can affect detritus processing, food web structure, and habitat condition[7, 8]. Moreover, because crayfish are sensitive to environmental disturbances such as acidification, metal contamination, and low pH[14], monitoring their populations could serve as an early warning for freshwater degradation, especially with the high number of visitors at the reserve possibly being a source of contamination and disturbances and with streams near trails being less monitored than Hertel Lake. Our findings lay the groundwork for future monitoring at Gault Nature Reserve and highlight the need to integrate aquatic invertebrates into research and conservation efforts.

## Conclusion

This study provides the first documented record, to our knowledge, of crayfish at the Gault Nature Reserve of McGill University, with two different species occupying different habitats: the Virile crayfish in Lake Hertel and the Appalachian Brook crayfish in a nearby stream. This species/habitat distinction is the key result of this study, as it indicates new, previously unrecorded freshwater biodiversity and raises interesting ecological questions about habitat preference and community structure at the reserve. Morphometric tests also indicated that there were size differences between the two groups, although due to the confounding of species and habitat, these differences should be considered more as species/site differences than habitat differences in isolation. The colouration analyses did not support the substrate–matching hypothesis as proposed and are considered preliminary because of issues with image standardization and referencing the colour of the substrate.

Future work should include better–controlled colour measurements, seasonal sampling, and both sex– and age–specific analyses. Future research should clearly identify life stage and either limit analysis to adults or include size/developmental stage as a covariate. In addition, a reserve–wide survey alongside DNA genetic identification should be done to rule out hybridization and explore the occurrence of crayfish in other water bodies. Doing so, the initial hypothesis of this paper could be tested by comparing morphology within the same species across different habitats to identify environment–specific traits and confirm whether other species of crayfish inhabit the reserve. The use of transplant experiments could also clarify whether observed patterns are species–specific or environmentally driven.

Our findings address a major research gap in the reserve's freshwater ecosystem and underscore that even very well monitored, protected, and studied areas such as the Gault Nature Reserve can harbor omissions in its biodiversity monitoring coverage. Crustaceans and other freshwater organisms are among the only taxonomic groups without a species list at the reserve. Unfortunately, freshwater ecosystems are often the ones suffering from inferior research and conservation attention[1, 3]. Freshwater ecosystems cover a small fraction of Earth's surface, yet they contain a disproportionately large share of global diversity and are especially fragile and vulnerable to contamination from chemicals and invasive species[1, 2]. Crayfish are important components in these ecosystems; they are keystone players that shape community dynamics and ecosystem function[7, 8]. Through their feeding, burrowing, and sediment–disturbing activities, they accel-

erate leaf litter decomposition, redistribute fine sediments, and modify habitat complexity[8, 9]. These changes ripple across trophic levels, affecting fishes, algae, invertebrates, and even terrestrial species. Changes in crayfish populations would alter nutrient flow and resilience of freshwater systems[8, 9], as well as make them more vulnerable to the introduction of invasive species such as the keystone invader Rusty Crayfish (*Faxonius rusticus*), which causes devastating damage to Canadian ecosystems[6]. Furthermore, they serve as bioindicators of water quality and habitat health, making them a key species to monitor as an indicator of broader ecological health important for conservation planning.

## References


1. Darwall, W. & Vié, J.-C. Identifying important sites for conservation of freshwater biodiversity: extending the species-based approach. *Fish. Manag. Ecol* **12,** 287–293 (2005). 10.1111/j.1365-2400.2005.00449.x
2. Allan, J. & Flecker, A. Biodiversity conservation in running waters. *BioScience* **43,** 32–43 (1993). 10.2307/1312104
3. Dudgeon, D. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biol. Rev* **81,** 163–182 (2006). 10.1017/S1464793105006950
4. Cooke, M. & Xia, L. Impacts of Land–Based Recreation on Water Quality. *Nat. Areas J* **40,** 179–188 (2020). 10.3375/043.040.0209
5. The Effects of Noise on Aquatic Life II. *Adv. Exp. Med. Biol* **875** (2016). 10.1007/978-1-4939-2981-8
6. Brown, N. & Therriault, T. The hidden risk of keystone invaders in Canada: a case study using nonindigenous crayfish. *Can. J. Fish. Aquat. Sci* **79,** 1479–1496 (2022). 10.1139/cjfas-2021-0245
7. Statzner, B., Peltret, O. & Tomanova, S. Crayfish as geomorphic agents and ecosystem engineers: effect of a biomass gradient on baseflow and flood-induced transport of gravel and sand in experimental streams. *Freshw. Biol* **48,** 147–163 (2003). 10.1046/j.1365-2427.2003.00984.x
8. Reynolds, J., Souty–Grosset, C. & Richardson, A. Ecological roles of crayfish in freshwater and terrestrial habitats. *Freshw. Crayfish* **19,** 197–218 (2013). 10.5869/fc.2013.v19-2.197
9. Creed, R. & Reed, J. Ecosystem engineering by crayfish in a headwater stream community. *J. North Am. Benthol. Soc* **23,** 224–236 (2004). 10.1899/08873593
10. Hamr, P. & Berrill, M. The life histories of north-temperate populations of the crayfish Cambarus robustus and Cambarus bartoni. *Can. J. Zool* **63,** 2313–2322 (1985). 10.1139/z85-343
11. Dubé, J. & Desroches, J.-F. *Les Écrevisses Du Québec* 66. 2007. `https://belsp.uqtr.ca/id/eprint/628`.
12. Perry, W. Effects of water velocity on the size and shape of rusty crayfish, Orconectes rusticus. *Freshw. Sci* **32,** 1398–1409 (2013). 10.1899/12-166.2
13. Hacker, S. & Madin, L. Why habitat architecture and color are important to shrimps living in pelagic Sargassum: use of camouflage and plant-part mimicry. *Mar. Ecol. Prog. Ser* **70,** 143–155 (1991). 10.3354/meps070143
14. Hay, M., Pawlik, J., Duffy, J. & Fenical, W. Seaweed-herbivore-predator interactions: host-plant specialization reduces predation on small herbivores. *Oecologia* **81,** 418–427 (1989). 10.1007/BF00377093
15. Sacchi, R., Cancian, S., Ghia, D., Fea, G. & Coladonato, A. Color variation in signal crayfish Pacifastacus leniusculus. *Curr. Zool* **67,** 35–43 (2021). 10.1093/cz/zoaa031
16. Graham, Z. Prevalence and potential evolutionary significance of color variants in freshwater crayfishes (Decapoda: Astacidea. *J. Crustac. Biol* **43, ruad054** (2023). 10.1093/jcbiol/ruad054
17. Mathews, L. Intrapopulation variability in coloration is associated with reproductive season in the crayfish Faxonius virilis. *Curr. Zool* **71,** 251–262 (2025). 10.1093/cz/zoae046
18. Palma, A. & Steneck, R. Does variable coloration in juvenile marine crabs reduce risk of visual predation? *Ecology* **82,** 2961–2967 (2001). 10.1890/0012-9658(2001)082
19. Hartzell, S. Ontogenetic color change in the crayfish Cambarus b. bartonii and Faxonius obscurus : a test of Ortmann's hypotheses. *Freshw. Crayfish* **23,** 59–63 (2017). 10.5869/fc.2017.v23-1.59
20. Kaldre, K., Haugjärv, K., Liiva, M. & Gross, R. The effect of two different feeds on growth, carapace colour, maturation and mortality in marbled crayfish (Procambarus fallax f. virginalis. *Aquac. Int* **23,** 185–194 (2015). 10.1007/s10499-014-9807-1
21. Jezerinac, R. & Thoma, R. An illustrated key to the Ohio Cambarus and Fallicambarus (Decapoda: Cambaridae) with comments and a new subspecies record. *Ohio J. Sci* **84,** 120–125 .
22. Jezerinac, R. & Thoma, R. Morphological variations of Cambarus (Cambarus) bartonii cavatus (Decapoda: Cambaridae) from Ohio, with a diagnosis of the Ohio form. *Ohio J. Sci* **85,** 131–134 (1985).
23. Thacker, R., Hazlett, B., Esman, L., Stafford, C. & Keller, T. Color morphs of the crayfish Orconectes virilis. *Am. Midl. Nat* **129,** 182–199 (1993). 10.2307/2426447
24. Massgov. *Appalachian Brook Crayfish* 2025. `https://www.mass.gov/info-details/appalachian-brook-crayfish`.
25. Crocker, D. The crayfishes of New York State (Decapoda, Astacidae. *N.Y. State Mus. Sci. Serv. Bull* **355,** 1–97 (1957).
26. St. *Lawrence Global Observatory. Biodiversity crustaceans: ECCC crayfish species* 2019. `https://slgo.ca/en/biodiversity-crustaceans-eccc-crayfish-species/`.
27. Rozansky, Z., Larson, E. & Taylor, C. Invasive virile crayfish (Faxonius virilis Hagen, 1870) hybridizes with native spothanded crayfish (Faxonius punctimanus Creaser, 1933) in the Current River watershed of Missouri, US. *Aquat. Invasions* **16,** 690–709 (2021). 10.3391/ai.2021.16.4.07

Submitted: 02/26/2025
Accepted: 02/22/2026
Published: 04/10/2026

Research Article

[1]IGFAE, Universidade de Santiago de Compostela, E-15782 Galicia, Spain

[2]Heidelberg University, Grabengasse 1, 69117 Heidelberg, Germany

Keywords

quantum mechanics, perturbation theory

Email Correspondence

juan.alvarez.ruiz@rai.usc.es
juan.alvarez_ruiz@stud.uniheidelberg.de
meijian.li@usc.es

https://doi.org/10.26443/msurj.v21i1.302



Juan Álvarez Ruiz[1,2], Meijian Li[1]

# Time-Independent Perturbation Theory: Degeneracy Lifted to the Second Order

## Abstract

Although the fundamental principles of time-independent perturbation theory are well-established and widely applied, the case of degeneracy lifted at second order is seldom addressed explicitly in standard texts. This work fills this gap by presenting a systematic derivation of time-independent perturbation theory that accounts for the added complexity introduced by non-trivial degeneracies. We provide a general procedure and corresponding formulae for calculating state and energy corrections to any order, with degeneracy lifted to the second order. Additionally, we apply these formulae to solve a prototype problem of a dressed quark system. This work not only advances the theoretical framework but also provides practical tools for tackling complex problems in quantum physics.

## Introduction

Time-independent perturbation theory, formulated within the Hamiltonian framework is a cornerstone of quantum mechanics. It has been widely used to address a variety of physical problems, ranging from atomic phenomena such as the Zeeman and Stark effects[1] to the analysis of chemical bonding and rotational–vibrational corrections in molecular physics[2]. The basic idea is to start with a simple system for which the exact solution is known and add a term representing a perturbation to the Hamiltonian. Assuming that the perturbation is small, the eigenvalues and eigenstates of the perturbed system can be expressed as asymptotic series, in which "corrections" to those of the unperturbed simple system are expanded order by order. The complicated system can therefore be studied based on knowledge of the simpler one. If the eigenvalues of the unperturbed system are degenerate, the arbitrariness in the choice of the basis within the degenerate subspace makes the naive first-order energy correction ill-defined. In such cases, standard perturbation theory must be reformulated to properly resolve the degeneracy.

It is worth noting that a variety of perturbative methods have been developed to address specific physical problems with degeneracies. In molecular spectroscopy, the contact transformation formalism—closely related to canonical Van Vleck perturbation theory[3,4]—has been widely used in the treatment of rotation–vibration energies of polyatomic molecules[5–10]. These approaches rely on near-identity unitary (canonical) transformations to eliminate interactions order by order in the perturbation parameter and to construct an effective Hamiltonian within a selected subspace. While the formalism can, in principle, be extended to arbitrary order within a recursive canonical-transformation framework, the algebraic complexity increases rapidly at higher orders. In practice, applications to vibrationally excited polyatomic molecules have been carried out up to fourth order in several studies[11,12].

In this work, we present a systematic and comprehensive derivation of time-independent perturbation theory, offering a unified treatment that extends existing works. We address the non-degenerate case, the degenerate case with degeneracy lifted to the first order, and, crucially, the degenerate case with degeneracy lifted to the second order. While time-independent perturbation theory is discussed in various quantum mechanics textbooks, providing methods for the first two cases[1,13] and introducing elementary theory for the degenerate cases[14], our derivations go further. Notably, prior work presents an insightful derivation for the second-order degenerate case under the special condition of completely unbroken first-order degeneracy and in the lowest orders[15]. Our work extends this by generalizing the second-order degenerate case to encompass a broader range of scenarios, providing a systematic and complete derivation that is valid for all orders, thereby contributing to a more comprehensive development of the subject.

## Perturbation Theory of a Non-Degenerate System

Suppose that the problem Hamiltonian can be written as a summation of an unperturbed Hamiltonian $H_0$ and a perturbation $\lambda\delta H$. The exact solution of the Time-Independent Schrödinger Equation (TISE) with $H_0$ is known,

$$H_0 \left|\psi_n^{(0)}\right\rangle = E_n^{(0)} \left|\psi_n^{(0)}\right\rangle \ , \quad n = 1, 2, \ldots, N \ , \tag{1}$$

where $\left|\psi_n^{(0)}\right\rangle$ the $n$-th normalized eigenstate with eigenvalue $E_n^{(0)}$, with $N$ the dimension of the problem space. Here, we consider a non-degenerate system, which means that

$$E_n^{(0)} \neq E_m^{(0)} \ , \quad \text{for } n \neq m \ . \tag{2}$$

The perturbed TISE for the $n$-th eigenstate becomes

$$(H_0 + \lambda\delta H) \left|\psi_n\right\rangle = E_n \left|\psi_n\right\rangle \ . \tag{3}$$

At $\lambda = 0$, we recover the unperturbed problem as in Equation 1. We can then expect the solution to the problem to depend on $\lambda$, with the limiting case recovering the original solution as $\lambda \to 0$.

We expand the eigenvalues and eigenstates as a power series in $\lambda^*$:

$$E_n = E_n^{(0)} + E_n^{(1)}\lambda + E_n^{(2)}\lambda^2 + \mathcal{O}(\lambda^3) = \sum_{i=0}^{\infty} E_n^{(i)}\lambda^i \ , \tag{4a}$$

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + |\psi_n^{(1)}\rangle\,\lambda + |\psi_n^{(2)}\rangle\,\lambda^2 + \mathcal{O}(\lambda^3) = \sum_{i=0}^{\infty} |\psi_n^{(i)}\rangle\,\lambda^i \ . \tag{4b}$$

Here and throughout the paper, we use the upper index with parenthesis "$(i)$" to indicate that the quantity is of the $i$-th order in $\lambda$. We impose a condition that all the state corrections contain no vector along the unperturbed state, without loss of generality[15],

$$\langle\psi_n^{(0)}|\psi_n^{(i)}\rangle = \delta_{i,0} \ , \quad i = 0, 1, 2, \dots \ , \tag{5}$$

and it follows from Equation 4b that $\langle\psi_n^{(0)}|\psi_n\rangle = 1$.

Putting Equations 4b and 4a back to Equation 3, we get

$$H_0\,|\psi_n^{(0)}\rangle + \sum_{i=1}^{\infty}\left[H_0\,|\psi_n^{(i)}\rangle + \delta H\,|\psi_n^{(i-1)}\rangle\right]\lambda^i = E_n^{(0)}\,|\psi_n^{(0)}\rangle + \sum_{i=1}^{\infty}\lambda^i\sum_{j=0}^{i} E_n^{(j)}\,|\psi_n^{(i-j)}\rangle \ . \tag{6}$$

The equation holds only if it is satisfied at each order of $\lambda$. Therefore, we have,

$$H_0\,|\psi_n^{(0)}\rangle = E_n^{(0)}\,|\psi_n^{(0)}\rangle \ , \tag{7}$$

at $i = 0$, and

$$H_0\,|\psi_n^{(i)}\rangle + \delta H\,|\psi_n^{(i-1)}\rangle - \sum_{j=0}^{i} E_n^{(j)}\,|\psi_n^{(i-j)}\rangle = 0 \ , \tag{8}$$

for $i = 1, 2, \dots$, from which one can solve for the perturbed wavefunction and energy.

### The First Order Correction

To obtain first order correction to the eigenvalue $E_n^{(1)}$, we project Equation 8 with index $i = 1$ on the eigenstate $\langle\psi_n^{(0)}|$,

$$\langle\psi_n^{(0)}|H_0|\psi_n^{(1)}\rangle + \langle\psi_n^{(0)}|\delta H|\psi_n^{(0)}\rangle - E_n^{(0)}\,\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle - E_n^{(1)}\,\langle\psi_n^{(0)}|\psi_n^{(0)}\rangle = 0 \ . \tag{9}$$

The first and third terms cancel out and we end up with,

$$E_n^{(1)} = \langle\psi_n^{(0)}|\delta H|\psi_n^{(0)}\rangle \ . \tag{10}$$

It means that the diagonal matrix elements of $\delta H$ gives the 1st order energy correction, as illustrated in Figure 1. Let us define $\delta H_{m,n} \equiv \langle\psi_m^{(0)}|\delta H|\psi_n^{(0)}\rangle$. To obtain the first order correction to the eigenstate, we write the first order correction in terms of the unperturbed basis states,

$$|\psi_n^{(1)}\rangle = \sum_m |\psi_m^{(0)}\rangle\,\langle\psi_m^{(0)}|\psi_n^{(1)}\rangle = \sum_{m\neq n} c_m^{(n)(1)}\,|\psi_m^{(0)}\rangle \ , \tag{11}$$

in which we have applied the condition in Equation 5. To find the coefficients $c_m^{(n)(1)}$, we project Equation 8 with index $i = 1$ on the eigenstate $\langle\psi_m^{(0)}|$ where $m \neq n$,

$$\langle\psi_m^{(0)}|H_0|\psi_n^{(1)}\rangle + \langle\psi_m^{(0)}|\delta H|\psi_n^{(0)}\rangle - E_n^{(0)}\,\langle\psi_m^{(0)}|\psi_n^{(1)}\rangle - E_n^{(1)}\,\langle\psi_m^{(0)}|\psi_n^{(0)}\rangle = 0 \ . \tag{12}$$

* Here and throughout the paper, we assume that the relevant eigenvalues and eigenstates admit an analytic continuation in $\lambda$ in a neighborhood of $\lambda = 0$. In exceptional cases (e.g., in the presence of intruder states), such a continuation may fail to exist.

$$H_0 = \begin{pmatrix} E_1^{(0)} & & & \\ & E_2^{(0)} & & \\ & & \ddots & \\ & & & E_N^{(0)} \end{pmatrix} ,$$

$$\delta H = \begin{pmatrix} E_1^{(1)} & & & \\ & E_2^{(1)} & & \\ & & \ddots & \\ & & & E_N^{(1)} \end{pmatrix}$$

**Figure 1.** Matrix representations of the Hamiltonian operator $H_0$ and $\delta H$ in the unperturbed basis for a non-degenerate system. The diagonal matrix elements of the perturbation $\delta H$ gives the first order energy correction.

The last term vanishes by orthogonality,

$$c_m^{(n)(1)}(E_m^{(0)} - E_n^{(0)}) = -\,\langle\psi_m^0|\delta H|\psi_n^0\rangle \, . \tag{13}$$

For non-degenerate systems, it follows that

$$c_m^{(n)(1)} = -\frac{\delta H_{m,n}}{E_m^{(0)} - E_n^{(0)}} \ . \tag{14}$$

### Higher Order Correction

We can generalize the aforementioned procedure for higher order corrections, for which we obtain a recursive relation. We first find the corrections to the energy by projecting Equation 8 to $|\psi_n^{(0)}\rangle$,

$$\langle\psi_n^{(0)}|H_0|\psi_n^{(i)}\rangle + \langle\psi_n^{(0)}|\delta H|\psi_n^{(i-1)}\rangle - \sum_{j=1}^{i-1} E_n^{(j)}\,\langle\psi_n^{(0)}|\psi_n^{(i-j)}\rangle - E_n^{(0)}\,\langle\psi_n^{(0)}|\psi_n^{(i)}\rangle = E_n^{(i)}\,\langle\psi_n^{(0)}|\psi_n^{(0)}\rangle \ , \tag{15}$$

The first and forth terms cancel out by seeing that $|\psi_n^{(0)}\rangle$ is the eigenstate of $H_0$ with eigenvalue $E_n^{(0)}$. The third term vanishes by orthogonality. We therefore have

$$E_n^{(i)} = \langle\psi_n^{(0)}|\delta H|\psi_n^{(i-1)}\rangle \, , \qquad i = 1, 2, \dots \ . \tag{16}$$

This means that the $i$-th order energy correction can be determined by knowing the $(i-1)$-th order state correction. In particular, the second order correction is

$$E_n^{(2)} = \sum_{m\neq n} \frac{|\delta H_{m,n}|^2}{E_n^{(0)} - E_m^{(0)}} . \tag{17}$$

Now, we expand the $i$-th order eigenstates in terms of the unperturbed basis states,

$$|\psi_n^{(i)}\rangle = \sum_m c_m^{(n)(i)}\,|\psi_m^{(0)}\rangle \ . \tag{18}$$

The orthonormal condition of the unperturbed states gives $c_m^{(n)(0)} = \delta_{m,n}$ and the condition of the state correction given by Equation 5 can be written as $c_n^{(n)(i)} = 0$ when $i \geq 1$.

We project Equation 8 on the eigenstate $\langle\psi_k^{(0)}|$ with $k \neq n$,

$$c_k^{(n)(i)}(E_k^{(0)} - E_n^{(0)}) + \sum_m c_m^{(n)(i-1)} \langle\psi_k^{(0)}|\delta H|\psi_m^{(0)}\rangle - \sum_{j=1}^{i-1} E_n^{(j)} c_k^{(n)(i-j)} = 0 \,. \quad (19)$$

The $j = i$ term vanishes due to state orthogonality. We have a recursive relation for the coefficient $c_k^{(n)(i)}$ at different order $i$. For a non-degenerate system, where $E_k^{(0)} \neq E_n^{(0)}$ for $k \neq n$, one can solve for $c_k^{(n)(i)}$ as

$$c_k^{(n)(i)} = \frac{1}{E_k^{(0)} - E_n^{(0)}} \left[ -\sum_m c_m^{(n)(i-1)} \delta H_{k,m} + \sum_{j=1}^{i-1} E_n^{(j)} c_k^{(n)(i-j)} \right] . \quad (20)$$

## Perturbation of a Degenerate System, Degeneracy Lifted to the First Order

Now we consider perturbation theory in a degenerate case, where for the initial Hamiltonian there are more than one eigenstates having the same eigenvalue. The procedure for non-degenerate systems as described in the first section no longer fully applies, as Equations 17 and 20 can become singular with $E_k^0 - E_n^0 = 0$ $(k \neq n)$ in the denominator.

Without loss of generality, suppose there is a $g$-fold degeneracy in the unperturbed Hamiltonian, i.e., there are $g$ eigenstates of $H_0$ with the same eigenvalue $E_D^{(0)\dagger}$. We denote these states by $\{|D; m^{(0)}\rangle\}$ with $m = 1, 2, ..., g$, and their spanned subspace by $D$. The Hilbert space can be written as $\mathcal{H} = D \oplus \mathbb{V}_\perp$, in which $\mathbb{V}_\perp$ is the non-degenerate subspace orthogonal to $D$ and one can calculate its perturbation by using the non-degenerate perturbation procedure as explained in the first section.

Note that due to the degeneracy, any linear combination of $\{|D; m^{(0)}\rangle\}$ is also an eigenstate of $H_0$ with eigenvalue $E_D^{(0)}$. To compute the state and energy correction in powers of $\lambda$, as in Equations 4a and 4b, we need to use the correct unperturbed state that the corresponding perturbed state approaches in the limit $\lambda \to 0$. Let us denote the states of the subspace $D$ after perturbation as $\{|D; n\rangle\}$ with $n = 1, 2, ..., g$, which in general do not belong to $D$ but in the limit of $\lambda \to 0$ become $\{|D; n^{(0)}\rangle\}$. For this reason, we will refer to the states $\{|D; n^{(0)}\rangle\}$ as the "good basis". The set $\{|D; n^{(0)}\rangle\}$ is a basis of $D$, which however are not necessarily the same with $\{|D; m^{(0)}\rangle\}$. As such, we have

$$|D; n\rangle = |D; n^{(0)}\rangle + |D; n^{(1)}\rangle \lambda + |D; n^{(2)}\rangle \lambda^2 + \mathcal{O}(\lambda^3) \,, \quad (21a)$$

$$E_{D;n} = E_D^{(0)} + E_{D;n}^{(1)} \lambda + E_{D;n}^{(2)} \lambda^2 + \mathcal{O}(\lambda^3) \,. \quad (21b)$$

We have imposed the orthonormal condition Equation 5 on $\{|D; n^{(i)}\rangle\}$, in the same manner as for $\{|\psi_n^{(i)}\rangle\}$ in the discussion of the non-degenerate system. Note that Equations 6-8, which are obtained by rearranging the Schrödinger equation order by order in $\lambda$, also apply to these states upon making the replacements $E_n \to E_{D;n}$, $E_n^{(i)} \to E_{D;n}^{(i)}$, $|\psi_n\rangle \to |D; n\rangle$, and $|\psi_n^{(i)}\rangle \to |D; n^{(i)}\rangle$.

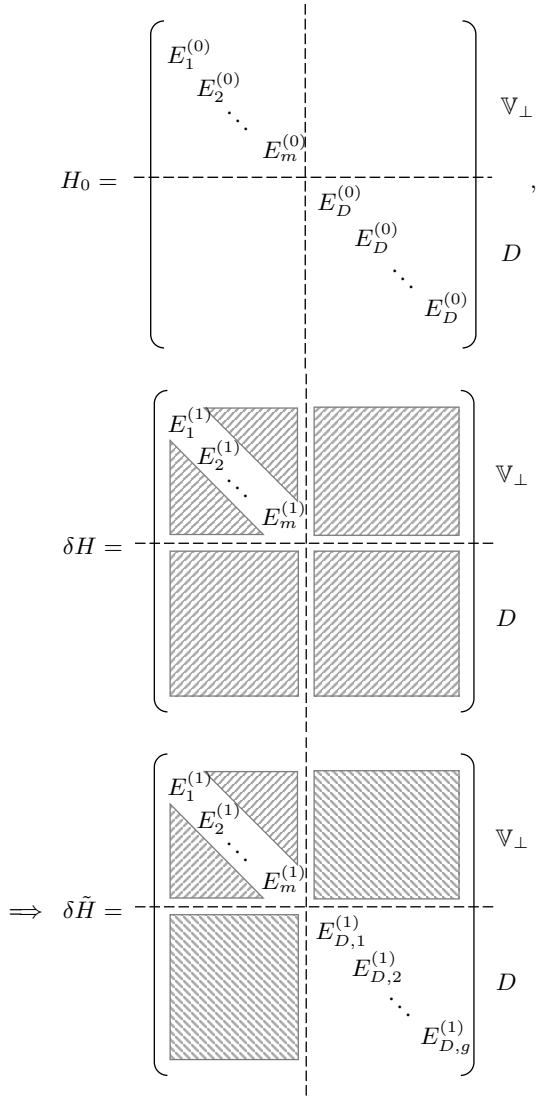



**Figure 2.** Matrix representations of the Hamiltonian operator for a system with degeneracy, with $\mathbb{V}_\perp$ denoting the non-degenerate subspace and $D$ the degenerate subspace. From top to bottom: $H_0$ in the unperturbed basis, $\delta H$ in the unperturbed basis, and $\delta\tilde{H}$, the matrix representation of $\delta H$ in the "good basis" where it is diagonal in $D$.

### First Order Equation

We project Equation 8 (with the replacement $E_n^{(i)} \to E_{D;n}^{(i)}$ and $|\psi_n^{(i)}\rangle \to |D; n^{(i)}\rangle$) for the case $i = 1$, onto $\langle D; l^{(0)}|$, an arbitrary state from the set $\{|D; n^{(0)}\rangle\}$ that we are looking for,

$$\langle D; l^{(0)}|H_0|D; n^{(1)}\rangle + \langle D; l^{(0)}|\delta H|D; n^{(0)}\rangle - E_D^{(0)} \langle D; l^{(0)}|D; n^{(1)}\rangle - E_{D;n}^{(1)} \langle D; l^{(0)}|D; n^{(0)}\rangle = 0 \,. \quad (22)$$

The first and third terms cancel out and we end up with

$$E_{D;n}^{(1)} \delta_{n,l} = \langle D; l^{(0)}|\delta H|D; n^{(0)}\rangle \,. \quad (23)$$

This means that the basis $\{|D; n^{(0)}\rangle\}$ diagonalizes $\delta H$ in $D$, and $E_{D;n}^{(1)}$s are the eigenvalues, as illustrated in Figure 2. For simplicity, let us define $\delta H_{D;l,D;n} \equiv \langle D; l^{(0)}|\delta H|D; n^{(0)}\rangle$, such that $E_{D;n}^{(1)} = \delta H_{D;n,D;n}$.

To continue, we project the same first order TISE, onto a state in the $\mathbb{V}_\perp$

[†] In cases with multiple degenerate subspaces, say $D_1$, $D_2$, . . ., each having a different eigenvalue $E_{D_1}^{(0)}$, $E_{D_2}^{(0)}$, . . ., one can apply the procedure described in this section to each subspace.

subspace, $\langle\psi_p^{(0)}|$,

$$\langle\psi_p^{(0)}|H_0|D;n^{(1)}\rangle + \langle\psi_p^{(0)}|\delta H|D;n^{(0)}\rangle - E_D^{(0)}\,\langle\psi_p^{(0)}|D;n^{(1)}\rangle - E_{D;n}^{(1)}\,\langle\psi_p^{(0)}|D;n^{(0)}\rangle = 0\ . \quad (24)$$

The last term vanishes due to orthogonality, and we are left with

$$\langle\psi_p^{(0)}|D;n^{(1)}\rangle = -\frac{\delta H_{p,D;n}}{E_p^{(0)} - E_D^{(0)}}\ , \quad (25)$$

in which we have defined $\delta H_{p,D;n} \equiv \langle\psi_p^{(0)}|\delta H|D;n^{(0)}\rangle$ for convenience. Note that we only get the part of the first order state in the $\mathbb{V}_\perp$ subspace, and that in $D$ is still missing,

$$|D;n^{(1)}\rangle = \sum_{p\notin D}|\psi_p^{(0)}\rangle\,\langle\psi_p^{(0)}|D;n^{(1)}\rangle + \sum_{l\in D,l\neq n}|D;l^{(0)}\rangle\,\langle D;l^{(0)}|D;n^{(1)}\rangle\ . \quad (26)$$

## Second Order Equation

We project the second order TISE, i.e., Equation 8 with index $i = 2$, onto $\langle D;l^{(0)}|$,

$$\langle D;l^{(0)}|(H_0 - E_D^{(0)})|D;n^{(2)}\rangle - \langle D;l^{(0)}|(E_{D;n}^{(1)} - \delta H)|D;n^{(1)}\rangle - E_{D;n}^{(2)}\,\langle D;l^{(0)}|D;n^{(0)}\rangle = 0\ . \quad (27)$$

The first term vanishes immediately. Substituting Equation 26 and performing some simplifications, we arrive at

$$(E_{D;n}^{(1)} - E_{D;l}^{(1)})\,\langle D;l^{(0)}|D;n^{(1)}\rangle - \sum_{p\notin D}\delta H_{D;l,p}\,\langle\psi_p^{(0)}|D;n^{(1)}\rangle + E_{D;n}^{(2)}\delta_{l,n} = 0\ . \quad (28)$$

We already know the second term from Equation 25. We take $l \neq n$ to extract the remaining information in $|D;n^{(1)}\rangle$. Here, we assume that the degeneracy is completely resolved so that $E_{D;n}^{(1)} \neq E_{D;l}^{(1)}$ when $l \neq n$. Therefore,

$$\langle D;l^{(0)}|D;n^{(1)}\rangle = \sum_{p\notin D} -\frac{\delta H_{D;l,p}\delta H_{p,D;n}}{(E_{D;n}^{(1)} - E_{D;l}^{(1)})(E_p^{(0)} - E_D^{(0)})}\ . \quad (29)$$

We obtain the first order state correction with Equations 25, 26, and 29. We take $l = n$ in Equation 28 and obtain the second order energy correction as

$$E_{D;n}^{(2)} = \sum_{p\notin D} -\frac{|\delta H_{p,D;n}|^2}{E_p^{(0)} - E_D^{(0)}}\ . \quad (30)$$

To complete the calculation at the second order equation, we project the second-order TISE onto a state in the $\mathbb{V}_\perp$ subspace, $\langle\psi_p^{(0)}|$,

$$\langle\psi_p^{(0)}|(H_0 - E_D^{(0)})|D;n^{(2)}\rangle - \langle\psi_p^{(0)}|(E_{D;n}^{(1)} - \delta H)|D;n^{(1)}\rangle - E_{D;n}^{(2)}\,\langle\psi_p^{(0)}|D;n^{(0)}\rangle = 0\ . \quad (31)$$

The last term vanishes due to state orthogonality, and the first order corrections are already known. We therefore get the second order state correction in the $\mathbb{V}_\perp$ subspace,

$$\langle\psi_p^{(0)}|D;n^{(2)}\rangle = \frac{1}{E_p^{(0)} - E_D^{(0)}}\Bigg[E_{D;n}^{(1)}\,\langle\psi_p^{(0)}|D;n^{(1)}\rangle - \sum_{q\notin D}\delta H_{p,q}\,\langle\psi_q^{(0)}|D;n^{(1)}\rangle - \sum_{l\in D,l\neq n}\delta H_{p,D;l}\,\langle D;l^{(0)}|D;n^{(1)}\rangle\Bigg]\ . \quad (32)$$

## Higher Order Corrections

Let us generalize the aforementioned procedure for higher-order corrections. Note that for states in the $\mathbb{V}_\perp$ subspace, the results shown in the first section apply directly. The $i$-th order eigenstates, expressed in terms of the unperturbed basis states, can be written as:

$$|D;n^{(i)}\rangle = (\mathrm{P}_{\mathbb{V}_\perp} + \mathrm{P}_D)\,|D;n^{(i)}\rangle = \sum_{p\notin D}|\psi_p^{(0)}\rangle\,c_p^{(D;n)(i)} + \sum_{l\in D}|D;l^{(0)}\rangle\,c_{D;l}^{(D;n)(i)}\ , \quad (33)$$

in which $c_p^{(D;n)(i)} \equiv \langle\psi_p^{(0)}|D;n^{(i)}\rangle$ and $c_{D;l}^{(D;n)(i)} \equiv \langle D;l^{(0)}|D;n^{(i)}\rangle$. The orthonormal condition of the unperturbed states gives $c_{D;l}^{(D;n)(0)} = \delta_{l,n}$ and the condition of the state correction given by Equation 5 can be written as $c_{D;n}^{(D;l)(i)} = 0$ for $l = n$ when $i \geq 1$.

From the $i$-th order TISE, one can express $E_{D;n}^{(i)}$, $c_p^{(D;n)(i)}$s and $c_{D;l}^{(D;n)(i-1)}$s (for $i \geq 2$ in the last case) in terms of lower-order results, as shown in the first and second order equations. We now proceed with the $i$-th order equation, given by Equation 8,

$$H_0\,|D;n^{(i)}\rangle + \delta H\,|D;n^{(i-1)}\rangle - \sum_{j=0}^{i}E_{D;n}^{(j)}\,|D;n^{(i-j)}\rangle = 0\ , \quad (34)$$

and continue from here.

- Projection onto $\mathbb{V}_\perp$
  Projecting Equation 34 onto a state in the $\mathbb{V}_\perp$ subspace, $\langle\psi_p^{(0)}|$, gives

  $$(E_p^{(0)} - E_D^{(0)})\,\langle\psi_p^{(0)}|D;n^{(i)}\rangle + \langle\psi_p^{(0)}|\delta H|D;n^{(i-1)}\rangle - \sum_{j=1}^{i-1}E_{D;n}^{(j)}\,\langle\psi_p^{(0)}|D;n^{(i-j)}\rangle = 0\ . \quad (35)$$

  Expanding the state correction with Equation 33, we get

  $$c_p^{(D;n)(i)} = -\frac{1}{E_p^{(0)} - E_D^{(0)}}\Bigg[\sum_{q\notin D}\delta H_{p,q}c_q^{(D;n)(i-1)} + \sum_{l\in D}\delta H_{p,D;l}c_{D;l}^{(D;n)(i-1)} - \sum_{j=1}^{i-1}E_{D;n}^{(j)}c_p^{(D;n)(i-j)}\Bigg]\ . \quad (36)$$

  All terms on the right-hand side are known from lower-order equations.

- Projection onto $D$
  Projecting Equation 34 onto a state in the $D$ subspace, $\langle D;l^{(0)}|$, gives

  $$\langle D;l^{(0)}|\delta H|D;n^{(i-1)}\rangle - \sum_{j=1}^{i-1}E_{D;n}^{(j)}\,\langle D;l^{(0)}|D;n^{(i-j)}\rangle - E_{D;n}^{(i)}\,\langle D;l^{(0)}|D;n^{(0)}\rangle = 0\ . \quad (37)$$

  In the case of $l = n$, we have

  $$E_{D;n}^{(i)} = \langle D;n^{(0)}|\delta H|D;n^{(i-1)}\rangle\ . \quad (38)$$

  In the case of $l \neq n$, we have

  $$c_{D;l}^{(D;n)(i-1)} = \frac{1}{E_{D;l}^{(1)} - E_{D;n}^{(1)}}\Bigg[\sum_{j=2}^{i-1}E_{D;n}^{(j)}c_{D;l}^{(D;n)(i-j)} - \sum_{p\notin D}\delta H_{D;l,p}c_p^{(D;n)(i-1)}\Bigg]\ . \quad (39)$$

We summarize the general procedure for the perturbative theory of a degenerate system whose degeneracy is resolved at first order as follows:

i) Identify the degenerate subspace $D$ in the basis of unperturbed eigenstates of $H_0$. Diagonalize $\delta H$ in this subspace, and use its eigenstates as the new unperturbed states, with the eigenvalues giving the first-order state correction.

ii) Calculate the perturbation correction to states whose unperturbed states are in the degenerate subspace using Equations 36, 38, and 39.

iii) Calculate the perturbation correction to states whose unperturbed states are in the non-degenerate subspace using Equations 16 and 20.

## Perturbation of a Degenerate System, Degeneracy Lifted to the Second Order

When the degeneracy is not fully resolved in first order, the method in the previous fails, as Equation 29 becomes singular. In this case, we cannot determine the correct zeroth-order basis states from the first-order perturbation alone. This section addresses the scenario where the degeneracy is partially or not resolved at first order but is completely resolved at second order.

We consider a system with a $g$-fold degeneracy on a subspace $D$. We carry on the setup that the full Hilbert space of the problem can be written as $\mathcal{H} = D \oplus V_\perp$. The subspace $D$ is spanned by the unperturbed eigenstates $\{|D; n^{(0)}\rangle\}$ $(n = 1, 2, \ldots, g)$ (with unperturbed eigenvalue $E_D^{(0)}$) that already diagonalize the perturbation $\delta H$ in $D$. Their first order corrections, as can be obtained by Equation 23, are $\{E_{D_1}^{(1)}, E_{D_2}^{(1)}, \ldots, E_{D_B}^{(1)}\}$ where $B < g$; note that in the scenario of $B = g$ the degeneracy is completely broken at first order and the problem is solved according to the previous procedure. We now express $D$ as a direct sum of the subspaces with different first order correction, $D = \bigoplus_{I=1}^{B} D_I$, where $D_I$ is the subspace spanned by the unperturbed eigenstates that have the same first order energy correction $E_{D_I}^{(1)}$. The dimensions of $D_I$ is $N_I$ such that $\sum_{I=1}^{B} N_I = g$. For convenience, let us denote the unperturbed basis states $\{|D; n^{(0)}\rangle\}$ that are in the $D_I$ subspace with the subscript $I$ as $\{|D_I; n_I^{(0)}\rangle\}$ $(n_I = 1, 2, \ldots, N_I)$. For a subspace $D_I$ with $N_I > 1$, any linear combination of $\{|D; n^{(0)}\rangle\}$ is still a $H_0$ eigenstate with eigenvalue $E_D^{(0)}$, and our goal here is to find the "good basis" that the corresponding perturbed state approaches in the limit $\lambda \to 0$.

Let us denote the "good basis" by $\{|D_I; \alpha_I^{(0)}\rangle\}$ $(\alpha_I = 1, 2, \ldots, N_I)$, for which we use Greek letters as the state index to distinguish those of the original basis. We can write the $\alpha_I$-th state of the $D_I$ subspace as a linear combination of the original basis states,

$$|D_I; \alpha_I^{(0)}\rangle = \sum_{n_I=1}^{N_I} v_{n_I}^{\alpha_I} |D_I; n_I^{(0)}\rangle \ . \tag{40}$$

Here, $v_{n_I}^{\alpha_I} \equiv \langle D_I; n_I^{(0)} | D_I; \alpha_I^{(0)}\rangle$. Note that $\delta H$ remains diagonal in $D$ in the "good basis". We can expand the corresponding perturbed state and energy in powers of $\lambda$,

$$\begin{aligned} |D_I; \alpha_I\rangle =& |D_I; \alpha_I^{(0)}\rangle + |D_I; \alpha_I^{(1)}\rangle \lambda + |D_I; \alpha_I^{(2)}\rangle \lambda^2 \\ &+ \mathcal{O}(\lambda^3) \ , \end{aligned} \tag{41a}$$

$$E_{D_I;\alpha_I} = E_D^{(0)} + E_{D_I}^{(1)} \lambda + E_{D_I;\alpha_I}^{(2)} \lambda^2 + \mathcal{O}(\lambda^3) \ . \tag{41b}$$

As before, we apply the orthogonality condition on $\{|D_I; \alpha_I^{(i)}\rangle\}$ as defined in Equation 5. The eigenvalue equation, as of Equation 8, for $|D_I; \alpha_I^{(i)}\rangle$, becomes

$$\mathcal{O}(1) : (H_0 - E_D^{(0)}) |D_I; \alpha_I^{(0)}\rangle = 0 \ , \tag{42a}$$

$$\mathcal{O}(\lambda) : (H_0 - E_D^{(0)}) |D_I; \alpha_I^{(1)}\rangle = (E_{D_I}^{(1)} - \delta H) |D_I; \alpha_I^{(0)}\rangle \ , \tag{42b}$$

$$\begin{aligned} \mathcal{O}(\lambda^2) : (H_0 - E_D^{(0)}) |D_I; \alpha_I^{(2)}\rangle =& (E_{D_I}^{(1)} - \delta H) |D_I; \alpha_I^{(1)}\rangle \\ &+ E_{D_I;\alpha_I}^{(2)} |D_I; \alpha_I^{(0)}\rangle \ , \end{aligned} \tag{42c}$$

$$\begin{aligned} \mathcal{O}(\lambda^3) : (H_0 - E_D^{(0)}) |D_I; \alpha_I^{(3)}\rangle =& (E_{D_I}^{(1)} - \delta H) |D_I; \alpha_I^{(2)}\rangle \\ &+ E_{D_I;\alpha_I}^{(2)} |D_I; \alpha_I^{(1)}\rangle + E_{D_I;\alpha_I}^{(3)} |D_I; \alpha_I^{(0)}\rangle \ , \end{aligned} \tag{42d}$$

$$\tag{42e}$$

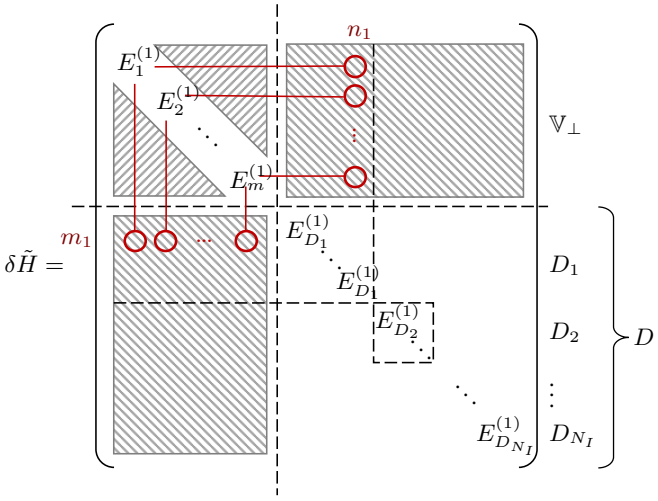



**Figure 3.** For a system with degeneracy lifted to the second order, diagonalizing the perturbation $\delta H$ in the $E^{(0)}$-degenerate subspace $D$ does not completely break the degeneracy (c.f. Figure 2), and we have $D = D_1 \oplus D_2 \oplus \ldots \oplus D_{N_I}$, where $D_I$'s are $E^{(1)}$-degenerate subspaces. The "good basis" diagonalizes $\mathbf{\Delta}_{D_I}^{(2)}$ as defined in Equation 51. The red circles denote the matrix elements involved in calculating $\Delta_{D_1}^{(2)}|_{m_1,n_1}$ as an example.

### First Order Equation

From the first order equation, we obtain the first order energy correction $E_{D_I}^{(1)} = \delta H_{D_I;\alpha_I,D_I;\alpha_I}$, similar to Equation 23, and the first order state correction in $\mathbb{V}_\perp$, analogous to Equation 25, which is:

$$\langle \psi_p^{(0)} | D_I; \alpha_I^{(1)}\rangle = -\frac{\delta H_{p,D_I;\alpha_I}}{E_p^{(0)} - E_D^{(0)}} = -\frac{\sum_{n_I=1}^{N_I} v_{n_I}^{\alpha_I} \delta H_{p,D_I;n_I}}{E_p^{(0)} - E_D^{(0)}} \ . \tag{43}$$

We have defined $\delta H_{p,D_I;\alpha_I} \equiv \langle \psi_p^{(0)} | \delta H | D_I; \alpha_I^{(0)}\rangle$ and similarly, $\delta H_{p,D_I;n_I} \equiv \langle \psi_p^{(0)} | \delta H | D_I; n_I^{(0)}\rangle$ for convenience. Note that $|D_I; n_I^{(0)}\rangle$ is just $|D; n^{(0)}\rangle$ in Equation 25 with the subspace index $I$ written out.

### Second Order Equation

In the second order equation, unlike before, we must account for the remaining degeneracy after the first order perturbation. Instead of Equa-

tion 26, we project $|D_I;\alpha_I^{(1)}\rangle$ onto three subspaces, $\mathbb{V}_\perp$, $D-D_I$, and $D_I$,

$$\begin{aligned}|D_I;\alpha_I^{(1)}\rangle &= (\mathrm{P}_{\mathbb{V}_\perp}+\mathrm{P}_{D-D_I}+\mathrm{P}_{D_I})\,|D_I;\alpha_I^{(1)}\rangle \\ &= \sum_{p\notin D}|\psi_p^{(0)}\rangle\,\langle\psi_p^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ &\quad+\sum_{J\neq I}\sum_{n_J=1}^{N_J}|D_J;n_J^{(0)}\rangle\,\langle D_J;n_J^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ &\quad+\sum_{n_I=1}^{N_I}|D_I;n_I^{(0)}\rangle\,\langle D_I;n_I^{(0)}|D_I;\alpha_I^{(1)}\rangle\ .\end{aligned} \tag{44}$$

The projection onto $D-D_I$ gives us part of the first order state correction. Written explicitly, projecting Equation 42c onto $\langle D_J;n_J^{(0)}|$ with $J\neq I$, we get

$$\begin{aligned}&\langle D_J;n_J^{(0)}|(H_0-E_D^{(0)})\,|D_I;\alpha_I^{(2)}\rangle \\ &\quad=\langle D_J;n_J^{(0)}|\,(E_{D_I}^{(1)}-\delta H)\,|D_I;\alpha_I^{(1)}\rangle \\ &\qquad+E_{D_I;\alpha_I}^{(2)}\,\langle D_J;n_J^{(0)}|D_I;\alpha_I^{(0)}\rangle\ .\end{aligned} \tag{45}$$

The first and second terms vanish, and by using Equation 44 we obtain

$$\begin{aligned}E_{D_I}^{(1)}\,\langle D_J;n_J^{(0)}|D_I;\alpha_I^{(1)}\rangle &= \sum_{p\notin D}\delta H_{D_J;n_J,p}\,\langle\psi_p^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ &\quad+\delta H_{D_J;n_J,D_J;n_J}\,\langle D_J;n_J^{(0)}|D_I;\alpha_I^{(1)}\rangle\ .\end{aligned} \tag{46}$$

Together with Equation 43, we have:

$$\langle D_J;n_J^{(0)}|D_I;\alpha_I^{(1)}\rangle = \sum_{p\notin D}-\frac{\delta H_{D_J;n_J,p}\delta H_{p,D_I;\alpha_I}}{\left(E_{D_I}^{(1)}-E_{D_J}^{(1)}\right)\left(E_p^{(0)}-E_D^{(0)}\right)}\ . \tag{47}$$

This result is similar to Equation 29, and recovers the latter when $N_I=1$.

Next, we project Equation 42c onto $D_I$,

$$\begin{aligned}&\langle D_I;n_I^{(0)}|(H_0-E_D^{(0)})\,|D_I;\alpha_I^{(2)}\rangle \\ &\quad=\langle D_I;n_I^{(0)}|\,(E_{D_I}^{(1)}-\delta H)\,|D_I;\alpha_I^{(1)}\rangle \\ &\qquad+E_{D_I;\alpha_I}^{(2)}\,\langle D_I;n_I^{(0)}|D_I;\alpha_I^{(0)}\rangle\ .\end{aligned} \tag{48}$$

The first term vanishes, and we recognize $v_{n_I}^{\alpha_I}$ in the last term. Applying Equation 44, we obtain:

$$\sum_{p\notin D}\delta H_{D_I;n_I,p}\,\langle\psi_p^{(0)}|D_I;\alpha_I^{(1)}\rangle - E_{D_I;\alpha_I}^{(2)}v_{n_I}^{\alpha_I}=0\ . \tag{49}$$

In the above equation, we used that that $\delta H$ is diagonal in $D$ and $E_{D_I}^{(1)}=\delta H_{D_I;n_I,D_I;n_I}$. Together with Equation 43, we have

$$-\sum_{p\notin D}\delta H_{D_I;n_I,p}\frac{\sum_{m_I=1}^{N_I}v_{m_I}^{\alpha_I}\delta H_{p,D_I;m_I}}{E_p^{(0)}-E_D^{(0)}}-E_{D_I;\alpha_I}^{(2)}v_{n_I}^{\alpha_I}=0\ . \tag{50}$$

We rewrite the above equation by defining the second-order perturbation operator $\mathbf{\Delta}_{D_I}^{(2)}$, with matrix elements

$$\begin{aligned}\Delta_{D_I}^{(2)}|_{n_I,m_I} &= \langle D_I;n_I^{(0)}|\mathbf{\Delta}_{D_I}^{(2)}|D_I;m_I^{(0)}\rangle \\ &\equiv -\sum_{p\notin D}\frac{\delta H_{D_I;n_I,p}\delta H_{p,D_I;m_I}}{E_p^{(0)}-E_D^{(0)}}\ .\end{aligned} \tag{51}$$

Physically, this is the energy fluctuation of an initial state in $D_I$ going through any intermediate states in $\mathbb{V}_\perp$ and back to $D_I$. It is structurally analogous to the second-order off-diagonal elements of the contact-transformed Hamiltonian in Darling-Dennison resonances in hydroxyl stretching polyads[16–18].

Equation 50 becomes

$$\sum_{m_I=1}^{N_I}\left[\Delta_{D_I}^{(2)}|_{n_I,m_I}-E_{D_I;\alpha_I}^{(2)}\delta_{n_I,m_I}\right]v_{m_I}^{\alpha_I}=0\ , \tag{52}$$

which is clearly an eigenvalue equation of vectors $\boldsymbol{v}^{\alpha_I}\equiv\{v_1^{\alpha_I},v_2^{\alpha_I},\ldots,v_{n_I}^{\alpha_I}\}$, with $\alpha_I=1,2,\ldots,N_I$. We therefore find the "good basis" by solving $\boldsymbol{v}^{\alpha_I}$ as the eigenvectors of $\mathbf{\Delta}_{D_I}^{(2)}$, and in the meantime, we obtain the second order energy correction $E_{D_I;\alpha_I}^{(2)}$ as the eigenvalues. We illustrate the relation of $\mathbf{\Delta}_{D_I}^{(2)}$ and the $\delta H$ matrix in Figure 3. Lastly, we project Equation 42c onto $\mathbb{V}_\perp$,

$$\begin{aligned}\langle\psi_p^{(0)}|\,(H_0-E_D^{(0)})\,|D_I;\alpha_I^{(2)}\rangle &= \langle\psi_p^{(0)}|\,(E_{D_I}^{(1)}-\delta H)\,|D_I;\alpha_I^{(1)}\rangle \\ &\quad+E_{D_I;\alpha_I}^{(2)}\,\langle\psi_p^{(0)}|D_I;\alpha_I^{(0)}\rangle\ .\end{aligned} \tag{53}$$

The last term vanishes by orthogonality. Decomposing the first order state by Equation 44, we get

$$\begin{aligned}\langle\psi_p^{(0)}|D_I;\alpha_I^{(2)}\rangle =&\frac{1}{E_p^{(0)}-E_D^{(0)}}\Bigg\{E_{D_I}^{(1)}\,\langle\psi_p^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ &-\sum_{q\notin D}\delta H_{p,q}\,\langle\psi_q^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ &-\sum_{J\neq I}\sum_{n_J=1}^{N_J}\delta H_{p,D_J;n_J}\,\langle D_J;n_J^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ &-\sum_{n_I=1}^{N_I}\delta H_{p,D_I;n_I}\,\langle D_I;n_I^{(0)}|D_I;\alpha_I^{(1)}\rangle\Bigg\}\ .\end{aligned} \tag{54}$$

The first three terms on the right-hand side of the equation are known from Equations 43 and 47. This gives a relation between two unknowns, but we have exhausted the second-order equation. Next, we proceed with the third order equation.

### Third Order Equation

We first project Equation 42d onto $D_I$, and now that we have found the "good basis", we use $\langle D_I;\beta_I^{(0)}|$. This gives us

$$\begin{aligned}&\langle D_I;\beta_I^{(0)}|(H_0-E_D^{(0)})\,|D_I;\alpha_I^{(3)}\rangle \\ &\quad=\langle D_I;\beta_I^{(0)}|\,(E_{D_I}^{(1)}-\delta H)\,|D_I;\alpha_I^{(2)}\rangle \\ &\qquad+E_{D_I;\alpha_I}^{(2)}\,\langle D_I;\beta_I^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ &\qquad+E_{D_I;\alpha_I}^{(3)}\,\langle D_I;\beta_I^{(0)}|D_I;\alpha_I^{(0)}\rangle\ .\end{aligned} \tag{55}$$

The terms on the left-hand side cancel out. In the case $\beta_I=\alpha_I$, we get, after applying the orthogonality condition of $|D_I;\alpha_I^{(i)}\rangle$ with respect to $|D_I;\alpha_I^{(0)}\rangle$ as defined in Equation 5,

$$\begin{aligned}E_{D_I;\alpha_I}^{(3)} &= \langle D_I;\alpha_I^{(0)}|\,\delta H\,|D_I;\alpha_I^{(2)}\rangle \\ &= \sum_{p\notin D}\delta H_{D_I;\alpha_I,p}\,\langle\psi_p^{(0)}|D_I;\alpha_I^{(2)}\rangle\ .\end{aligned} \tag{56}$$

We get to the second equation by noting that $\delta H$ is diagonal in $D$.

In the case $\beta_I\neq\alpha_I$, the last term of Equation 55 vanishes by state orthogonality, and we have

$$\begin{aligned}0 = \langle D_I;\beta_I^{(0)}|\,(E_{D_I}^{(1)}-\delta H)\,|D_I;\alpha_I^{(2)}\rangle \\ +E_{D_I;\alpha_I}^{(2)}\,\langle D_I;\beta_I^{(0)}|D_I;\alpha_I^{(1)}\rangle\ .\end{aligned} \tag{57}$$

We decompose the second-order state correction in the above equation as

$$|D_I;\alpha_I^{(2)}\rangle = \sum_{p\notin D}|\psi_p^{(0)}\rangle\langle\psi_p^{(0)}|D_I;\alpha_I^{(2)}\rangle + \sum_{J\neq I}\sum_{\gamma_J=1}^{N_J}|D_J;\alpha_J^{(0)}\rangle\langle D_J;\alpha_J^{(0)}|D_I;\alpha_I^{(2)}\rangle + \sum_{\gamma_I\neq\alpha_I}|D_I;\gamma_I^{(0)}\rangle\langle D_I;\gamma_I^{(0)}|D_I;\alpha_I^{(2)}\rangle \ . \quad (58)$$

The first component will survive in Equation 57; the second component will not contribute due to state orthogonality and $\delta H$ being diagonal in $D$; the third component will not contribute by seeing $\delta H \to E^{(1)}_{D_I}$ in $D_I$. As a result, we have

$$\sum_{p\notin D}\delta H_{D_I;\beta_I,p}\langle\psi_p^{(0)}|D_I;\alpha_I^{(2)}\rangle = E^{(2)}_{D_I;\alpha_I}\langle D_I;\beta_I^{(0)}|D_I;\alpha_I^{(1)}\rangle \ . \quad (59)$$

Combining with Equation 54, we get

$$\begin{aligned}E^{(2)}_{D_I;\alpha_I}\langle D_I;\beta_I^{(0)}|D_I;\alpha_I^{(1)}\rangle = \sum_{p\notin D}\delta H_{D_I;\beta_I,p}\frac{1}{E_p^{(0)}-E_D^{(0)}} \times\Bigg\{E^{(1)}_{D_I}\langle\psi_p^{(0)}|D_I;\alpha_I^{(1)}\rangle - \sum_{q\notin D}\delta H_{p,q}\langle\psi_q^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ - \sum_{J\neq I}\sum_{n_J=1}^{N_J}\delta H_{p,D_J;n_J}\langle D_J;n_J^{(0)}|D_I;\alpha_I^{(1)}\rangle \\ - \sum_{\gamma_I\neq\alpha_I}\delta H_{p,D_I;\gamma_I}\langle D_I;\gamma_I^{(0)}|D_I;\alpha_I^{(1)}\rangle\Bigg\} .\end{aligned} \quad (60)$$

Note that in writing the last term, we have switched the summation to the "good basis". Recall that we already know the first three terms on the right-hand side by Equations 43 and 47. Let us examine the last term,

$$-\sum_{p\notin D}\frac{\delta H_{D_I;\beta_I,p}\delta H_{p,D_I;\gamma_I}}{E_p^{(0)}-E_D^{(0)}} = \langle D_I;\beta_I^{(0)}|\mathbf{\Delta}^{(2)}_{D_I}|D_I;\gamma_I^{(0)}\rangle = E^{(2)}_{D_I;\beta_I}\delta_{\beta_I,\gamma_I} \ , \quad (61)$$

in which we recognize the $\mathbf{\Delta}^{(2)}_{D_I}$ operator, as in Equation 51, and recall that it is diagonal in the "good basis". By assuming that the degeneracy is completely resolved at second order, i.e., $E^{(2)}_{D_I;\alpha_I} \neq E^{(2)}_{D_I;\beta_I}$ for $\alpha_I \neq \beta_I$, we can write Equation 60 as

$$\begin{aligned}\langle D_I;\beta_I^{(0)}|D_I;\alpha_I^{(1)}\rangle = \frac{1}{E^{(2)}_{D_I;\alpha_I}-E^{(2)}_{D_I;\beta_I}}\sum_{p\notin D}\frac{\delta H_{D_I;\beta_I,p}}{E_p^{(0)}-E_D^{(0)}} \times\Bigg\{-E^{(1)}_{D_I}\frac{\delta H_{p,D_I;\alpha_I}}{E_p^{(0)}-E_D^{(0)}} + \sum_{q\notin D}\frac{\delta H_{p,q}\delta H_{q,D_I;\alpha_I}}{E_q^{(0)}-E_D^{(0)}} \\ + \sum_{J\neq I}\sum_{n_J=1}^{N_J}\sum_{q\notin D}\frac{\delta H_{p,D_J;n_J}\delta H_{D_J;n_J,q}\delta H_{q,D_I;\alpha_I}}{\left(E^{(1)}_{D_I}-E^{(1)}_{D_J}\right)\left(E_q^{(0)}-E_D^{(0)}\right)}\Bigg\} .\end{aligned} \quad (62)$$

Up to this point, we complete the derivation of the first order state correction. One can write out the full expression by putting Equations 43, 44, 47, and 62 together. It is then straightforward to get the second order state correction that comes from outside $D$ (i.e., $\mathbb{V}_\perp$) as in Equation 54 and the third order energy correction as in Equation 56.

We next project Equation 42d onto $D - D_I$, using the "good basis" $\langle D_J;\beta_J^{(0)}|\,(J\neq I)$,

$$\langle D_J;\beta_J^{(0)}|\,(H_0-E_D^{(0)})\,|D_I;\alpha_I^{(3)}\rangle = \langle D_J;\beta_J^{(0)}|\,(E^{(1)}_{D_I}-\delta H)\,|D_I;\alpha_I^{(2)}\rangle + E^{(2)}_{D_I;\alpha_I}\langle D_J;\beta_J^{(0)}|D_I;\alpha_I^{(1)}\rangle + E^{(3)}_{D_I;\alpha_I}\langle D_J;\beta_J^{(0)}|D_I;\alpha_I^{(0)}\rangle \ . \quad (63)$$

The terms on the left-hand side cancel out, and the last term on the right vanishes due to orthogonality. Therefore,

$$0 = \langle D_J;\beta_J^{(0)}|\,(E^{(1)}_{D_I}-\delta H)\,|D_I;\alpha_I^{(2)}\rangle + E^{(2)}_{D_I;\alpha_I}\langle D_J;\beta_J^{(0)}|D_I;\alpha_I^{(1)}\rangle \ . \quad (64)$$

By the state decomposition given in Equation 58, we obtain a second component of the second order state correction

$$\langle D_J;\beta_J^{(0)}|D_I;\alpha_I^{(2)}\rangle = \frac{-1}{E^{(1)}_{D_I}-E^{(1)}_{D_J}}\Bigg\{E^{(2)}_{D_I;\alpha_I}\langle D_J;\beta_J^{(0)}|D_I;\alpha_I^{(1)}\rangle - \sum_{p\notin D}\delta H_{D_J;\beta_J,p}\langle\psi_p^{(0)}|D_I;\alpha_I^{(2)}\rangle\Bigg\} . \quad (65)$$

Finally, we project Equation 42d onto $\mathbb{V}_\perp$ with $\langle\psi_p^{(0)}|$,

$$\langle\psi_p^{(0)}|\,(H_0-E_D^{(0)})\,|D_I;\alpha_I^{(3)}\rangle = \langle\psi_p^{(0)}|\,(E^{(1)}_{D_I}-\delta H)\,|D_I;\alpha_I^{(2)}\rangle + E^{(2)}_{D_I;\alpha_I}\langle\psi_p^{(0)}|D_I;\alpha_I^{(1)}\rangle + E^{(3)}_{D_I;\alpha_I}\langle\psi_p^{(0)}|D_I;\alpha_I^{(0)}\rangle \ , \quad (66)$$

and arrives at

$$\begin{aligned}\langle\psi_p^{(0)}|D_I;\alpha_I^{(3)}\rangle = \frac{1}{E_p^{(0)}-E_D^{(0)}}\Bigg\{E^{(1)}_{D_I}\langle\psi_p^{(0)}|D_I;\alpha_I^{(2)}\rangle \\ + E^{(2)}_{D_I}\langle\psi_p^{(0)}|D_I;\alpha_I^{(1)}\rangle - \sum_{q\notin D}\delta H_{p,q}\langle\psi_q^{(0)}|D_I;\alpha_I^{(2)}\rangle \\ - \sum_{J\neq I}\sum_{\beta_J=1}^{N_J}\delta H_{p,D_J;\beta_J}\langle D_J;\beta_J^{(0)}|D_I;\alpha_I^{(2)}\rangle \\ - \sum_{\gamma_I\neq\alpha_I}\delta H_{p,D_I;\gamma_I}\langle D_I;\gamma_I^{(0)}|D_I;\alpha_I^{(2)}\rangle\Bigg\} .\end{aligned} \quad (67)$$

This equation contains two unknowns: the left-hand side and the last term on the right-hand side. Since the third-order equation is exhausted, we need to go to the next order to solve for them.

## Higher Order Corrections

We can generalize the aforementioned procedure for higher order corrections, for which we will obtain a recursive relation of the corrections. Note that for states in the non-degenerate subspace, the formulas obtained before apply directly, and here we make the derivation for the states in the degenerate subspace $D$. The $i$-th order eigenstates in terms of the unperturbed basis states can be written as,

$$\begin{aligned}|D_I;\alpha_I^{(i)}\rangle &= (\mathrm{P}_{\mathbb{V}_\perp}+\mathrm{P}_{D-D_I}+\mathrm{P}_{D_I})\,|D_I;\alpha_I^{(i)}\rangle \\ &= \sum_{p\notin D}|\psi_p^{(0)}\rangle\, c_p^{(D_I;\alpha_I)(i)} \\ &\quad + \sum_{J\neq I}\sum_{\beta_J=1}^{N_J}|D_J;\beta_J^{(0)}\rangle\, c^{(D_I;\alpha_I)(i)}_{D_J;\beta_J} \\ &\quad + \sum_{\gamma_I=1}^{N_I}|D_I;\gamma_I^{(0)}\rangle\, c^{(D_I;\alpha_I)(i)}_{D_I;\gamma_I} \ .\end{aligned} \quad (68)$$

in which $c_\psi^{(D_I;\alpha_I)(i)} \equiv \langle\psi|D_I;\alpha_I^{(i)}\rangle$. From the $i$-th order TISE, we can get $E^{(i)}_{D_I;\alpha_I}$, $c_p^{(D_I;\alpha_I)(i-1)}$, $c^{(D_I;\alpha_I)(i-2)}_{D_I;\beta_I}(\beta_I \neq \alpha_I)$, $c^{(D_I;\alpha_I)(i-1)}_{D_J;\beta_J}(I \neq J)$ in terms of lower order results, and a relation between $c_p^{(D_I;\alpha_I)(i)}$ and $c^{(D_I;\alpha_I)(i-1)}_{D_I;\beta_I}(\beta_I \neq \alpha_I)$ [which will be solved with the $(i+1)$-th order

equation]. We have shown this with the eigenvalue equations up to the third order. We proceed with the $i(\geq 3)$-th order equation,

$$H_0 \left|D_I;\alpha_I^{(i)}\right\rangle + \delta H \left|D_I;\alpha_I^{(i-1)}\right\rangle - \sum_{j=0}^{i} E_{D_I;\alpha_I}^{(j)} \left|D_I;\alpha_I^{(i-j)}\right\rangle = 0 \ . \quad (69)$$

- Projection onto $D_I$
  Projecting Equation 69 onto a state in the $D_I$ subspace, $\langle D_I;\gamma_I^{(0)}|$, gives

$$\begin{aligned}\langle D_I;\gamma_I^{(0)}|\delta H|D_I;\alpha_I^{(i-1)}\rangle - \sum_{j=1}^{i-1} E_{D_I;\alpha_I}^{(j)} \langle D_I;\gamma_I^{(0)}|D_I;\alpha_I^{(i-j)}\rangle \\ - E_{D_I;\alpha_I}^{(i)} \langle D_I;\gamma_I^{(0)}|D_I;\alpha_I^{(0)}\rangle = 0 \ . \quad (70)\end{aligned}$$

  In the case of $\gamma_I = \alpha_I$, we have

$$E_{D_I;\alpha_I}^{(i)} = \langle D_I;\alpha_I^{(0)}|\delta H|D_I;\alpha_I^{(i-1)}\rangle \ . \quad (71)$$

  In the case of $\gamma_I \neq \alpha_I$, after expanding the state correction with Equation 68, we get

$$\begin{aligned}\sum_{p\notin D} \delta H_{D_I;\gamma_I,p} c_p^{(D_I;\alpha_I)(i-1)} - E_{D_I;\alpha_I}^{(2)} c_{D_I;\gamma_I}^{(D_I;\alpha_I)(i-2)} \\ - \sum_{j=3}^{i-1} E_{D_I;\alpha_I}^{(j)} c_{D_I;\gamma_I}^{(D_I;\alpha_I)(i-j)} = 0 \ . \quad (72)\end{aligned}$$

  The unknown terms at the current order are $c_p^{(D_I;\alpha_I)(i-1)}$ and $c_{D_I;\gamma_I}^{(D_I;\alpha_I)(i-2)}$. We will solve it in combination with the $(i-1)$-th order Equation 74

- Projection onto $\mathbb{V}_\perp$
  Projecting Equation 69 onto $\mathbb{V}_\perp$ gives

$$\begin{aligned}(E_p^{(0)} - E_D^{(0)}) \langle \psi_p^{(0)}|D_I;\alpha_I^{(i)}\rangle + \langle \psi_p^{(0)}|\delta H|D_I;\alpha_I^{(i-1)}\rangle \\ - \sum_{j=1}^{i-1} E_{D_I;\alpha_I}^{(j)} \langle \psi_p^{(0)}|D_I;\alpha_I^{(i-j)}\rangle = 0 \ , \quad (73)\end{aligned}$$

  Expanding the state correction with Equation 68, we get

$$\begin{aligned}c_p^{(D_I;\alpha_I)(i)} = -\frac{1}{E_p^{(0)} - E_D^{(0)}} \Bigg[ \sum_{q\notin D} \delta H_{p,q} c_q^{(D_I;\alpha_I)(i-1)} \\ + \sum_{J\neq I} \sum_{\beta_J=1}^{N_J} \delta H_{p,D_J;\beta_J} c_{D_J;\beta_J}^{(D_I;\alpha_I)(i-1)} \\ + \sum_{\gamma_I=1}^{N_I} \delta H_{p,D_I;\gamma_I} c_{D_I;\gamma_I}^{(D_I;\alpha_I)(i-1)} \\ - \sum_{j=1}^{i-1} E_{D_I;\alpha_I}^{(j)} c_p^{(D_I;\alpha_I)(i-j)} \Bigg] \ . \end{aligned} \quad (74)$$

  All terms except $c_p^{(D_I;\alpha_I)(i)}$ and $c_{D_I;\beta_I}^{(D_I;\gamma_I)(i-1)}$ are known from lower order equations. These two terms will be solved together with Equation 72 of the $(i+1)$-th order TISE.

- Combination with the $(i-1)$-th order equation
  Plugging the $(i-1)$-th order of Equation 74 into the first term of Equation 72 (which is obtained at the $i$-th order) gives

$$\begin{aligned}\sum_{p\notin D} \frac{\delta H_{D_I;\gamma_I,p}}{E_p^{(0)} - E_D^{(0)}} \Bigg[ \sum_{q\notin D} \delta H_{p,q} c_q^{(D_I;\alpha_I)(i-2)} \\ + \sum_{J\neq I} \sum_{\beta_J=1}^{N_J} \delta H_{p,D_J;\beta_J} c_{D_J;\beta_J}^{(D_I;\alpha_I)(i-2)} \\ + \sum_{\rho_I=1}^{N_I} \delta H_{p,D_I;\rho_I} c_{D_I;\rho_I}^{(D_I;\alpha_I)(i-2)} \\ - \sum_{j=1}^{i-2} E_{D_I;\alpha_I}^{(j)} c_p^{(D_I;\alpha_I)(i-1-j)} \Bigg] + E_{D_I;\alpha_I}^{(2)} c_{D_I;\gamma_I}^{(D_I;\alpha_I)(i-2)} \\ + \sum_{j=3}^{i-1} E_{D_I;\alpha_I}^{(j)} c_{D_I;\gamma_I}^{(D_I;\alpha_I)(i-j)} = 0 \ . \end{aligned} \quad (75)$$

  We simplify third term in the bracket by recognizing the $\mathbf{\Delta}_{D_I}^{(2)}$ operator, as in Equation 61, and obtain

$$\begin{aligned}c_{D_I;\gamma_I}^{(D_I;\alpha_I)(i-2)} = \frac{1}{E_{D_I;\gamma_I}^{(2)} - E_{D_I;\alpha_I}^{(2)}} \Bigg\{ \sum_{p\notin D} \frac{\delta H_{D_I;\gamma_I,p}}{E_p^{(0)} - E_D^{(0)}} \\ \times \Bigg[ \sum_{q\notin D} \delta H_{p,q} c_q^{(D_I;\alpha_I)(i-2)} \\ + \sum_{J\neq I} \sum_{\beta_J=1}^{N_J} \delta H_{p,D_J;\beta_J} c_{D_J;\beta_J}^{(D_I;\alpha_I)(i-2)} \\ - \sum_{j=1}^{i-2} E_{D_I;\alpha_I}^{(j)} c_p^{(D_I;\alpha_I)(i-1-j)} \Bigg] \\ + \sum_{j=3}^{i-1} E_{D_I;\alpha_I}^{(j)} c_{D_I;\gamma_I}^{(D_I;\alpha_I)(i-j)} \Bigg\} \ . \end{aligned} \quad (76)$$

  Substituting into Equation 74, we obtain $c_p^{(D_I;\alpha_I)(i-1)}$.

- Projection onto $D_J (J \neq I)$
  Projecting Equation 69 onto a state in the $D_J$ subspace, $\langle D_J;\beta_J^{(0)}|$, gives

$$\begin{aligned}\langle D_J;\beta_J^{(0)}|\delta H|D_I;\alpha_I^{(i-1)}\rangle \\ = \sum_{j=1}^{i-1} E_{D_I;\alpha_I}^{(j)} \langle D_J;\beta_J^{(0)}|D_I;\alpha_I^{(i-j)}\rangle \ , \quad (77)\end{aligned}$$

  Expanding the state correction with Equation 68, we get

$$\begin{aligned}c_{D_J;\beta_J}^{(D_I;\alpha_I)(i-1)} = \frac{1}{E_{D_I}^{(1)} - E_{D_J}^{(1)}} \\ \times \Bigg[ \sum_{p\notin D} \delta H_{D_J;\beta_J,p} c_p^{(D_I;\alpha_I)(i-1)} \\ - \sum_{j=2}^{i-1} E_{D_I;\alpha_I}^{(j)} c_{D_J;\beta_J}^{(D_I;\alpha_I)(i-j)} \Bigg] \end{aligned} \quad (78)$$

We summarize the general procedure for the perturbative theory of a degenerate system with degeneracy resolved at second order:

i) Identify the degenerate subspace $D$ in the basis of unperturbed eigenstates of $H_0$. Diagonalize $\delta H$ in $D$, and take the eigenstates as the trial unperturbed states in this subspace.

ii) Identify the remaining degeneracy in the trial basis space by writing $D = D_1 \oplus D_2 \oplus \cdots$, such that different $D_I$s have different $\delta H$ eigenvalues and each $D_I$ subspace is degenerate in $\delta H$.

iii) In each $D_I$ subspace, construct the second-order perturbation matrix $\mathbf{\Delta}^{(2)}_{D_I}$, according to Equation 51. Then solve its secular equation, Equation 52, by diagonalization and take the eigenstates as the correct unperturbed states.

iv) Calculate the perturbation correction to the states whose unperturbed states are in the degenerate subspace $D$ according to Equations 71, 74, 76, and 78.

v) Calculate the perturbation correction to the states whose unperturbed states are in the non-degenerate subspace according to Equations 16 and 20.

If the degeneracy is not fully removed in second order, one has to proceed to higher orders. In some cases, as in the problem discussed in the next section, the degeneracy may persist in all orders. In such scenarios, symmetries of the Hamiltonian can be used to decompose the Hilbert space into decoupled subspaces, where the degeneracy exists between subspaces, not within them. Perturbation theory can then be applied to each subspace separately.

## Application: Dressed Quark Prototype

We now apply time-independent perturbation theory to the dressed quark in Quantum Chromodynamics (QCD). In the QCD Hamiltonian formalism, a dressed quark is represented by the eigenstate in the expanded Fock space $|q\rangle + |qg\rangle + \cdots$ with the quantum numbers of a single quark. Here, we extract a prototype from the more extensively studied dressed quark scattering problem in high-energy nuclear physics[19].

Consider a simplified Hilbert space of $|q\rangle + |qg\rangle$, where the $|q\rangle$ sector contains the two spin states of the quark, $s_q = \pm 1/2$, and the $|qg\rangle$ sector contains the four quark-gluon spin states, $\{s_q, s_g\} = \{+1/2, +1\}, \{+1/2, -1\}, \{-1/2, +1\}$, and $\{-1/2, -1\}$. All other quantum numbers, including color and momentum (total for quark-gluon states), are chosen to be the same, allowing quark and quark-gluon states to couple through the $q \leftrightarrow qg$ vertex interaction. The resulting Hamiltonian is given by $H = H_0 + \lambda \delta H$, with

$$H_0 = \text{diag}\{a, a, b, b, b, b\} \,,$$
$$\delta H = \begin{bmatrix} 0 & 0 & u^R & v^R & w & 0 \\ 0 & 0 & 0 & -w & v^L & u^L \\ u^L & 0 & 0 & 0 & 0 & 0 \\ v^L & -w & 0 & 0 & 0 & 0 \\ w & v^R & 0 & 0 & 0 & 0 \\ 0 & u^R & 0 & 0 & 0 & 0 \end{bmatrix} . \tag{79}$$

The unperturbed Hamiltonian $H_0$ is degenerate in the $|q\rangle$ and the $|qg\rangle$ sectors, with $a, b > 0$, denoting the kinetic energy of the quark and the quark-gluon states respectively. The perturbation $\delta H$ is the interaction for the gluon emission and absorption, and its structure is extracted from the vertex interaction such that $u^{R/L} = ue^{\pm i\theta_u}$, $v^{R/L} = ve^{\pm i\theta_v}$, and $u, v, w > 0$.

The unperturbed eigenstates and eigenvalues are,

$$\begin{aligned} E^{(0)}_{1,m} = a \,, \ (m = 1, 2), \qquad & |v^{(0)}_{1,1}\rangle = [1,0,0,0,0,0]^T , \\ & |v^{(0)}_{1,2}\rangle = [0,1,0,0,0,0]^T , \\ E^{(0)}_{2,p} = b \,, \ (p = 1, 2, 3, 4), \qquad & |v^{(0)}_{2,1}\rangle = [0,0,1,0,0,0]^T , \\ & |v^{(0)}_{2,2}\rangle = [0,0,0,1,0,0]^T , \\ & |v^{(0)}_{2,3}\rangle = [0,0,0,0,1,0]^T , \\ & |v^{(0)}_{2,4}\rangle = [0,0,0,0,0,1]^T . \end{aligned} \tag{80}$$

There are two degenerate subspaces: $\mathbb{V}_1$ spanned by $v^{(0)}_{1,m}$, and $\mathbb{V}_2$ spanned by $v^{(0)}_{2,p}$. Since $\delta H$ is already diagonal within each subspace, the degeneracy remains unbroken at first order and we need to treat the degeneracy at second order. We write out the second-order energy fluctuation operator according to Equation 51. For the subspace $\mathbb{V}_1$,

$$\mathbf{\Delta}^{(2)}_{\mathbb{V}_1} = -\frac{1}{b-a} \begin{bmatrix} \Delta^2 & 0 \\ 0 & \Delta^2 \end{bmatrix} , \tag{81}$$

where we define $\Delta^2 \equiv u^2 + v^2 + w^2$. It is already diagonal, which means that the original basis is already the good basis in $\mathbb{V}_1$, and the diagonal elements are the second order energy correction,

$$\begin{aligned} & E^{(2)}_{1,1} = E^{(2)}_{1,2} = -\frac{1}{b-a}\Delta^2 , \\ & \{|\psi^{(0)}_{1,1}\rangle , |\psi^{(0)}_{1,2}\rangle\} = \{|v^{(0)}_{1,1}\rangle , |v^{(0)}_{1,2}\rangle\} . \end{aligned} \tag{82}$$

For the subspace $\mathbb{V}_2$,

$$\mathbf{\Delta}^{(2)}_{\mathbb{V}_2} = \frac{1}{b-a} \begin{bmatrix} u^2 & u^L v^R & u^L w & 0 \\ u^R v^L & v^2 + w^2 & 0 & -u^L w \\ u^R w & 0 & v^2 + w^2 & u^L v^R \\ 0 & -u^R w & u^R v^L & u^2 \end{bmatrix} . \tag{83}$$

By diagonalizing $\mathbf{\Delta}^{(2)}_{\mathbb{V}_2}$, we find the good basis and the second-order energy correction simultaneously,

$$\begin{aligned} E^{(2)}_{2,1} = 0, \qquad & |\psi^{(0)}_{2,1}\rangle = \frac{1}{\Delta u}[0,0,-u^L v^R, u^2, 0, w u^R]^T , \\ E^{(2)}_{2,2} = 0, \qquad & |\psi^{(0)}_{2,2}\rangle = \frac{1}{\Delta u}[0,0,-w u^L, 0, u^2, -v^L u^R]^T , \\ E^{(2)}_{2,3} = \frac{\Delta^2}{b-a}, \qquad & |\psi^{(0)}_{2,3}\rangle = \frac{1}{\Delta}[0,0,u^L, v^L, w, 0]^T , \\ E^{(2)}_{2,4} = \frac{\Delta^2}{b-a}, \qquad & |\psi^{(0)}_{2,4}\rangle = \frac{1}{\Delta}[0,0,0,-w, v^R, u^R]^T . \end{aligned} \tag{84}$$

Note that the degeneracy is not completely broken at second order. Let us write out the Hamiltonian in the good basis $\{|\psi^{(0)}_{n,m}\rangle\}$,

$$\tilde{H}_0 = \text{diag}\{a, a, b, b, b, b\} ,$$
$$\delta\tilde{H} = \begin{bmatrix} 0 & 0 & 0 & 0 & \Delta & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta & 0 & 0 & 0 & 0 \end{bmatrix} . \tag{85}$$

The system decouple into three blocks as $\tilde{H}_0 = \tilde{H}_{0,1} \oplus \tilde{H}_{0,2} \oplus \tilde{H}_{0,3}$ and $\delta\tilde{H} = \delta\tilde{H}_1 \oplus \delta\tilde{H}_2 \oplus \delta\tilde{H}_3$, with

$$\begin{aligned}
\mathbb{V}_{Q,1} &= \{|\psi^{(0)}_{1,1}\rangle, |\psi^{(0)}_{2,3}\rangle\}, \quad \tilde{H}_{0,1} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \delta\tilde{H}_1 = \begin{bmatrix} 0 & \Delta \\ \Delta & 0 \end{bmatrix}, \\
\mathbb{V}_{Q,2} &= \{|\psi^{(0)}_{1,2}\rangle, |\psi^{(0)}_{2,4}\rangle\}, \quad \tilde{H}_{0,2} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \delta\tilde{H}_2 = \begin{bmatrix} 0 & \Delta \\ \Delta & 0 \end{bmatrix}, \\
\mathbb{V}_{qge} &= \{|\psi^{(0)}_{2,1}\rangle, |\psi^{(0)}_{2,2}\rangle\}, \quad \tilde{H}_{0,3} = \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix}, \delta\tilde{H}_3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned} \tag{86}$$

We proceed the perturbation calculations separately for each block. The subspaces $\mathbb{V}_{Q,1}$ and $\mathbb{V}_{Q,2}$ have identical structures, reflecting the degeneracy of spin-up and spin-down dressed quarks. Applying nondegenerate perturbation theory via Equations 16 and 20, we obtain the solution for each subspace. For $\mathbb{V}_{Q,1}$,

$$\begin{aligned}
E_{1,1}(\lambda) &= a + \lambda^2 \frac{\Delta^2}{a-b} + \lambda^4 \frac{\Delta^4}{(b-a)^3} - \lambda^5 \frac{2\Delta^6}{(b-a)^5} + O(\lambda^8), \\
|\psi_{1,1}(\lambda)\rangle &= |\psi^{(0)}_{1,1}\rangle - \lambda \frac{\Delta}{b-a} |\psi^{(0)}_{2,3}\rangle + \lambda^3 \frac{\Delta^3}{(b-a)^3} |\psi^{(0)}_{2,3}\rangle \\
&\quad - \lambda^5 \frac{2\Delta^5}{(b-a)^5} |\psi^{(0)}_{2,3}\rangle + O(\lambda^7) , \\
E_{2,3}(\lambda) &= b + \lambda^2 \frac{\Delta^2}{b-a} + \lambda^4 \frac{\Delta^4}{(a-b)^3} - \lambda^5 \frac{2\Delta^6}{(a-b)^5} + O(\lambda^8), \\
|\psi_{2,3}(\lambda)\rangle &= |\psi^{(0)}_{2,3}\rangle - \lambda \frac{\Delta}{a-b} |\psi^{(0)}_{1,1}\rangle + \lambda^3 \frac{\Delta^3}{(a-b)^3} |\psi^{(0)}_{1,1}\rangle \\
&\quad - \lambda^5 \frac{2\Delta^5}{(a-b)^5} |\psi^{(0)}_{1,1}\rangle + O(\lambda^7) .
\end{aligned} \tag{87}$$

The result is the same for $\mathbb{V}_{Q,2}$ if we swap indices "1,1"→"1,2" and "2,3"→"2,4". The subspace $\mathbb{V}_{qge}$ is perturbation-free, so the unperturbed eigenstates are the exact solution,

$$\begin{aligned}
E_{2,1}(\lambda) &= b, \qquad |\psi_{2,1}(\lambda)\rangle = |\psi^{(0)}_{2,1}\rangle , \\
E_{2,2}(\lambda) &= b, \qquad |\psi_{2,2}(\lambda)\rangle = |\psi^{(0)}_{2,2}\rangle .
\end{aligned} \tag{88}$$

For comparison, we find the exact solutions of the full Hamiltonian $H$ by diagonalization. The eigenvector space splits into three degenerate subspaces with the following eigenvalues and eigenstates:

$$\begin{aligned}
E_1 &= \frac{1}{2}(a+b-\eta), \\
|\phi_{1,1}\rangle &= \frac{1}{N_-}[b-a+\eta, 0, -2\lambda u^L, -2\lambda v^L, -2\lambda w, 0]^T, \\
|\phi_{1,2}\rangle &= \frac{1}{N_-}[0, b-a+\eta, 0, 2\lambda w, -2\lambda v^R, -2\lambda u^R]^T, \\
E_2 &= \frac{1}{2}(a+b+\eta), \\
|\phi_{2,1}\rangle &= \frac{1}{N_+}[a-b+\eta, 0, 2\lambda u^L, 2\lambda v^L, 2\lambda w, 0]^T, \\
|\phi_{2,2}\rangle &= \frac{1}{N_+}[0, a-b+\eta, 0, -2\lambda w, 2\lambda v^R, 2\lambda u^R]^T, \\
E_3 &= b, \\
|\phi_{3,1}\rangle &= \frac{1}{\Delta u}[0, 0, -u^L v^R, u^2, 0, w u^R]^T, \\
|\phi_{3,2}\rangle &= \frac{1}{\Delta u}[0, 0, -w u^L, 0, u^2, -v^L u^R]^T.
\end{aligned} \tag{89}$$

Here, we define the quantities $\eta \equiv \sqrt{(a-b)^2 + 4\lambda^2\Delta^2}$ and $N_\pm \equiv \sqrt{(a-b\pm\eta)^2 + 4\lambda^2\Delta^2}$ to reduce the clutter in the algebra. In comparison with the perturbative results we obtained: $\phi_{1,1}$ and $\phi_{1,2}$ are $\psi_{1,1}$ and $\psi_{1,2}$ upon normalization, $\phi_{2,1}$ and $\phi_{2,2}$ are $\psi_{2,3}$ and $\psi_{2,4}$ upon normalization, and $\phi_{3,1}$ and $\phi_{3,2}$ are just $\psi_{2,1}$ and $\psi_{2,2}$. The eigenvalues also agree, seeing that $E_1 = E_{1,1} = E_{1,2}$, $E_2 = E_{2,3} = E_{2,4}$ and $E_3 = E_{2,1} = E_{2,2}$. Although diagonalizing the full Hamiltonian is feasible here due to the relatively small problem size, for higher-dimensional cases—such as those involving multiple momentum and color modes—perturbation theory offers valuable insights to reduce complexity and simplify calculations.

In this problem, the full Hamiltonian remains degenerate at all perturbation orders. Such degeneracy implies the existence of a symmetry operator **S** that commutes with $H$. Consequently, the Hamiltonian becomes block-diagonal in the **S** eigenspace, allowing each block to be solved independently. Here, the **S** operator is essentially the quark spin projection operator (containing the relative orbital angular momentum between the quark and the gluon). By diagonalizing $\mathbf{\Delta}^2$ using the second-order degenerate perturbation theory, we effectively block-diagonalize the Hamiltonian $\tilde{H}$.

## Conclusion

In this work, we have presented a systematic derivation of time-independent perturbation theory for systems in which degeneracy is lifted at second order. By carefully analyzing the structure of the perturbative expansion, we established a general procedure for computing energy and state corrections to arbitrary order within this framework. The resulting formulae clarify how higher-order effects generate effective couplings within an initially degenerate subspace and provide a transparent interpretation of the associated second-order operator.

When degeneracy persists at first order but is lifted at second order, the perturbative structure acquires additional subtleties that are not fully addressed in standard treatments. By organizing the expansion consistently and making explicit the role of intermediate states outside the degenerate subspace, we obtain expressions that are systematic and directly applicable to practical calculations. Beyond the prototype dressed quark example considered here, this formalism may apply to a wide range of degenerate or near-degenerate systems encountered in molecular spectroscopy, atomic and nuclear structure, condensed matter physics, and effective Hamiltonian approaches in quantum field theory. As an illustration, in the hydrogen atom the unperturbed energies depend only on the principal quantum number $n$; a perturbation preserving rotational symmetry at first order may split the $l$ degeneracy while leaving $m$ degenerate, with a second-order symmetry-breaking term lifting the remaining $m$ degeneracy—precisely the type of scenario addressed here.

We hope that the formulation presented here will serve both as a pedagogical clarification and as a practical tool for tackling complex perturbative problems in quantum physics.

## Acknowledgements

We thank T. Lappi, W. Qian, C. Salgado, and X. Zhao for valuable discussions. J.A is supported by IGFAE Summer Fellowships 2023 funded through the Collaboration Agreement between the Regional Ministry of Culture, Education and University and the USC for the development of strategic R&D&I actions. M.L. is supported by Xunta de Galicia (CIGUS accr editation), European Union ERDF, the Spanish Research State Agency under project PID2020-119632GB-I00, and European Research Council under project ERC-2018-ADG-835105 YoctoLHC.

## References


1. Sakurai, J. J. & Napolitano, J. *Modern Quantum Mechanics* ISBN: 978-1-108-58728-0 (Cambridge University Press, Oct. 2020).
2. Atkins, P. W. & Friedman, R. S. *Molecular Quantum Mechanics* 5th edn. ISBN: 9780199697403 (Oxford University Press, Oxford, UK, 2011).
3. Van Vleck, J. H. On $\sigma$-Type Doubling and Electron Spin in the Spectra of Diatomic Molecules. *Phys. Rev.* **33,** 467–506 (1929). https://doi.org/10.1103/PhysRev.33.467
4. Kemble, E. C. *The Fundamental Principles of Quantum Mechanics : With Elementary Applications* (McGraw-Hill, New York, 1937).
5. Niroomand-Rad, A. & Parker, P. M. Sequential contact transformation formulation of asymmetric-rotator vibration-rotation Hamiltonian. *Journal of Molecular Spectroscopy* **85,** 40–54 (1981). https://doi.org/10.1016/0022-2852(81)90308-8
6. Goldsmith, M., Amat, G. & Nielsen, H. H. Higher Order Rotation-Vibration Energies of Polyatomic Molecules. I. *The Journal of Chemical Physics* **24,** 1178–1182 (1956). https://doi.org/10.1063/1.1742736
7. Amat, G., Goldsmith, M. & Nielsen, H. H. Higher Order Rotation-Vibration Energies of Polyatomic Molecules. II. *The Journal of Chemical Physics* **27,** 838–844 (1957). https://doi.org/10.1063/1.1743862
8. Amat, G. & Nielsen, H. H. Higher Order Rotation-Vibration Energies of Polyatomic Molecules. III. *The Journal of Chemical Physics* **27,** 845–850 (1957). https://doi.org/10.1063/1.1743863
9. Amat, G. & Nielsen, H. H. Higher Order Rotation-Vibration Energies of Polyatomic Molecules. IV. *The Journal of Chemical Physics* **29,** 665–672 (1958). https://doi.org/10.1063/1.1744557
10. Amat, G. & Nielsen, H. H. Higher-Order Rotation-Vibration Energies of Polyatomic Molecules. V. *The Journal of Chemical Physics* **36,** 1859–1865 (1962). https://doi.org/10.1063/1.1701280
11. Sibert Edwin L., I. Theoretical studies of vibrationally excited polyatomic molecules using canonical Van Vleck perturbation theory. *The Journal of Chemical Physics* **88,** 4378–4390 (1988). https://doi.org/10.1063/1.453797
12. Krasnoshchekov, S. V., Isayeva, E. V. & Stepanov, N. F. Numerical-Analytic Implementation of the Higher-Order Canonical Van Vleck Perturbation Theory for the Interpretation of Medium-Sized Molecule Vibrational Spectra. *The Journal of Physical Chemistry A* **116,** 3691–3709 (2012). https://doi.org/10.1021/jp211400w
13. Griffiths, D. J. & Schroeter, D. F. *Introduction to Quantum Mechanics* 3rd edn (Cambridge University Press, 2018).
14. Messiah, A. *Quantum Mechanics Quantum Mechanics* **v. 2.** ISBN: 9780720400458 (Elsevier Science, 1961).
15. Zwiebach, B. *Mastering Quantum Mechanics: Essentials, Theory, and Applications* ISBN: 9780262046138 (MIT Press, 2022).
16. Darling, B. T. & Dennison, D. M. The Water Vapor Molecule. *Phys. Rev.* **57,** 128–139 (1940). https://doi.org/10.1103/PhysRev.57.128
17. Lehmann, K. K. Beyond the x-K relations. *Molecular Physics* **66,** 1129–1137 (1989). https://doi.org/10.1080/00268978900100751
18. Matthews, D. A., Vázquez, J. & Stanton, J. F. Calculated stretching overtone levels and Darling–Dennison resonances in water: a triumph of simple theoretical approaches. *Molecular Physics* **105,** 2659–2666 (2007). https://doi.org/10.1080/00268970701618424
19. Li, M., Lappi, T., Zhao, X. & Salgado, C. A. Scattering and gluon emission of physical quarks in a SU(3) colored field. *Phys. Rev. D* **112,** 016009 (2025). https://doi.org/10.1103/x99p-tfwg

Submitted: 01/16/2026
Accepted: 02/21/2026
Published: 04/10/2026

Research Article

[1] Earth Systems Science, McGill University

Keywords

Boreal peatlands, Net ecosystem exchange, Conservation prioritization, Human footprint

Email Correspondence

sophie.piret@mail.mcgill.ca

https://doi.org/10.26443/msurj.v21i1.412



Piret, Sophie J.[1], Saul, Quinn[1], Châtelier, Simone[1]

# Spatial Modeling of Canadian Boreal Peatland Carbon Sinks: An Integrative Framework to Support Climate and Development Policy

## Abstract

Peatlands are among the most effective carbon sinks, yet their stability is increasingly threatened by expanding land use and infrastructure. Spatially explicit information remains limited on where Canada's boreal peatlands combine high $CO_2$ sink strength with low human pressure. We mapped growing-season $CO_2$ sink potential and identified low-disturbance areas across Canadian boreal peatlands by integrating climate and human footprint data. Using AmeriFlux eddy-covariance data from four Canadian boreal peatland flux-tower sites from 2007–2019, we modeled climatic controls on net ecosystem exchange (NEE) of $CO_2$ and upscaled predictions with ERA5-Land reanalysis data. A pooled multiple linear regression using air temperature, subsurface soil temperature, and incoming shortwave radiation explained 57% of NEE variance (adjusted $R^2$=0.57). Comparisons of tower-measured and ERA5-Land climatic variables showed strong agreement for shortwave radiation and modest temperature biases, highlighting uncertainty in regional upscaling. Applying the model across boreal peatlands predicted the strongest sinks in the Northwest Territories and weaker sinks toward the northern boreal margin and mountainous regions. We then derived Predicted Carbon Sink Strength (PCSS) and combined it with an inverted human footprint to construct a Boreal Peatland Conservation Index (BPCI) to prioritize areas with high sink strength and low human footprint. High BPCI values were concentrated in the Hudson Bay Lowlands and the Northwest Territories. Because predictions are based on four flux-tower sites and represent growing-season $CO_2$ exchange, results are best interpreted as screening-level guidance. Overall, this integrative framework provides an interpretable, updatable approach for identifying boreal peatland regions where protecting $CO_2$ sinks may deliver high climate benefits with low development conflict.

## Introduction

Peatlands are wetlands formed by the accumulation of partially decomposed organic material under water-saturated, low-oxygen conditions. Globally, peatlands cover only ~3% of Earth's land surface yet store ~30% of the world's soil carbon[1, 2]. Total peatland carbon stocks are estimated at 500–600 Pg C, exceeding the carbon stored in the world's forests[2]. Their importance is especially pronounced in Canada, the location of roughly a quarter of the world's peatlands[3], primarily in the boreal region[4], underscoring Canada's role in climate-mitigation planning.

Beyond carbon storage, peatlands support biodiversity, regulate hydrology, improve water quality, and hold cultural significance for many Indigenous communities[1]. Yet affected by drainage, agriculture, peat extraction, mining, and infrastructure development[1, 2]. Drainage increases oxygen exposure and accelerates decomposition, shifting peatlands from long-term carbon sinks toward persistent carbon sources[2]. Wildlife Conservation Society Canada reports that only 13% of Canada's peatlands were protected as of 2023[5], leaving ~87% vulnerable to industrial and land-use pressures. Conservation planning requires maps that extend peatland $CO_2$ sink estimates beyond the limited network of field sites and place sink potential in the context of human disturbance.

Peatland carbon dynamics are commonly assessed using net ecosystem exchange (NEE), which quantifies the net $CO_2$ flux between an ecosystem and the atmosphere[6]. NEE reflects the balance between $CO_2$ uptake through photosynthesis, quantified by gross primary production (GPP), and $CO_2$ emissions via ecosystem respiration (R)[7, 8]. NEE is commonly expressed as $NEE = R - GPP$, where negative NEE values indicate a net carbon sink, and positive values indicate a net carbon source[9]. Because photosynthesis and respiration are reduced under freezing conditions, peatland $CO_2$ exchange is most concentrated during the biologically active season[10].

Climate regulates peatland NEE through its effects on photosynthesis and respiration. Air and soil temperatures (TA, TS) affect both processes, with warming often increasing ecosystem respiration and potentially reducing net $CO_2$ uptake[9, 11]. Incoming shortwave radiation (SWin) provides energy for photosynthesis and is commonly associated with stronger $CO_2$ uptake during the growing season[7, 9]. Moisture also shapes NEE because saturation suppresses decomposition, whereas drying can increase aerobic respiration and constrain plant productivity. Although water-table depth is a key hydrologic control on peatland carbon exchange[8, 11], it is often challenging to represent at regional scales. This creates an opportunity to use spatially continuous climatic predictors to upscale peatland $CO_2$ exchange beyond point measurements.

The Canadian Model for Peatlands developed by Bona et al. depends on multiple simplifying assumptions about complex processes such as temperature-sensitive decay and water-table dynamics[11]. This motivates simpler regression-based approaches that leverage widely available climatic variables to predict growing-season NEE. Prior work suggests that wetland NEE can often be captured with relatively simple climate–flux

relationships, motivating regression-based upscaling from AmeriFlux flux-tower observations[9]. While successfully applied in coastal wetlands[7], a comparable regression-based framework to predict and map NEE across boreal Canadian peatlands has not yet been developed. Moreover, the integration of carbon-sink maps with human disturbance metrics remains understudied. This need is especially acute as Canada expands northern access for critical minerals, increasing land-use pressures in regions such as the Hudson Bay Lowlands and the Ring of Fire.

To address this gap, we developed a regression model calibrated with AmeriFlux NEE tower observations and applied it to ERA5-Land climate data to predict NEE across boreal Canada. We used these predictions to construct a Canadian Boreal Peatland Conservation Index (BPCI) to support conservation and climate-mitigation planning. The BPCI integrates predicted NEE with a human footprint index that captures human disturbance from built infrastructure, transportation networks, and land-use conversion[12]. By combining a climate-driven indicator of sink strength with human footprint data, the BPCI highlights regions for further assessment where conservation may maximize $CO_2$-uptake benefits while prioritizing areas with lower disturbance.

Specifically, we address the following questions:

i) Which widely available climatic predictors are most informative for modeling growing-season NEE at boreal peatland AmeriFlux sites?
ii) How does predicted growing-season NEE vary across boreal Canadian peatlands based on key climatic predictors?
iii) How closely do ERA5-derived climatic predictors match tower-observed values during the growing season? What does this imply for upscaling uncertainty?
iv) Where should conservation be prioritized to maximize predicted $CO_2$ sink strength while minimizing human disturbance?

## Methods

### Study area and sites

Our study focuses on the four available boreal Canadian peatland sites in the AmeriFlux eddy-covariance network: the Attawapiskat River Bog (CA-ARB), Attawapiskat River Fen (CA-ARF), Churchill Fen site 1 (CA-CF1), and Scotty Creek Bog (CA-SCB). The combined dataset spans from 2007 to 2019, with individual sites providing varying record lengths. Sites include two bogs and two fens and span latitudes from 52.695°N to 61.308°N. Across sites, mean annual precipitation ranges from 388 to 700 mm, mean annual temperature ranges from −6.5 to −1.3 °C, and elevations range from 16.5 to 280 m (Table 1).

Global Lakes and Wetlands Database v2 (GLWD v2) provides percent coverage per grid cell for 33 wetland and waterbody types from 1990–2020 at 15 arc-seconds (~500 m) resolution. The broader study area of Canadian boreal peatlands was derived from the GLWD v2 by extracting the Arctic/Boreal peatland classes, both forested and unforested[13]. We then constrained the layer to the Canadian boreal extent in ArcGIS Pro to avoid conflating peatlands with different climatic regimes and ecological controls[14]. This produced a boreal peatland extent and fractional-cover layer used for spatial upscaling and index calculations.

### Datasets and climate predictors

AmeriFlux is a network of sites that measures ecosystem-atmosphere fluxes using eddy-covariance (EC) towers, along with associated meteorological variables[15]. We used AmeriFlux NEE and meteorological products processed with standard EC corrections and quality control, including low-turbulence filtering and marginal distribution sampling (MDS) gap-filling[16, 17]. GPP and ecosystem respiration are then partitioned from NEE using established flux-partitioning algorithms[16, 18].

ERA5-Land is a global, hourly reanalysis dataset at a resolution of 0.1° (~9 km) produced using a blend of observations and physics-based weather models[19].

We used the 2020 Global 100 m Terrestrial Human Footprint (HFP-100) dataset, which provides global maps of human pressure to quantify anthropogenic disturbance across the study region[12].

We initially considered climate variables shared between AmeriFlux and ERA5-Land based on previous studies[9, 20]. Predictors included 2 m air temperature (TA; °C), soil temperature (TS; °C), incoming shortwave radiation (SWin; W $m^2$), incoming longwave radiation (LWin; W $m^2$), and precipitation (P; mm).

### Preprocessing and growing-season definition

We standardized AmeriFlux missing values and screened candidate predictors for collinearity prior to regression. Predictor screening based on pairwise Spearman correlations and variance inflation factor (VIF) diagnostics is described in Supplementary Materials (Figure S2). Based on this screening, we retained TA, TS, and SWin for regression and excluded LWin due to strong collinearity with TA.

AmeriFlux soil temperature was harmonized to the ERA5-Land 28–100 cm layer (layer 3) to support consistent upscaling. Details on sensor-depth handling are provided in Supplementary Materials.

To focus on biologically relevant periods, months with mean monthly TA<0°C were excluded from analysis. The rule was applied consistently to tower observations and ERA5-Land predictors. This ensured datasets reflected only the growing season, focusing the analysis on months when $CO_2$ flux magnitudes are highest in these systems (Supplementary Figure S1). ERA5-Land variables were converted to monthly mean values and AmeriFlux units before agreement testing, growing-season definition, and upscaling. All subsequent analyses used the pooled AmeriFlux dataset.

### Exploratory analysis and AmeriFlux regression

We used monthly mean NEE (μmol $CO_2$ $m^{-2}$ $s^{-1}$; negative values indicate net $CO_2$ uptake) as the response variable. We analyzed monthly means to align temporal scale across sites and ERA5-Land, reduce short-term noise and gap-filling artifacts, and focus on seasonal-scale controls suitable for regional screening. To explore broad relationships between climatic variables and NEE, and to inform predictor selection, we computed monthly Spearman's rank correlation coefficients (ρ) between each climatic predictor and NEE (α=0.05), a rank-based measure robust to non-normality and outliers[21]. We also visually assessed scatterplots between NEE and climatic variables to examine linearity, variance structure, and cross-site consistency prior to modeling (Figure 2). Climatic variables included TA, TS, SWin, and log-transformed precipitation (log(P+1)) as a proxy for moisture conditions, noting that precipitation does not fully represent water-table dynamics. Precipitation showed no significant association with NEE in the correlation analysis and was therefore not retained in the final regression model; see results for further discussion of variable selection.

Multiple linear regression was implemented using ordinary least squares to quantify the effects of selected climatic variables on NEE. The final regression for spatial upscaling used TA, TS, and SWin. Regression coefficients were evaluated at α=0.05, and model performance was assessed using

adjusted $R^2$ and root mean square error.

### ERA5-Land agreement and spatial upscaling

We evaluated agreement between tower observations and their equivalent ERA5-Land predictors to assess potential reanalysis biases relevant to NEE upscaling[19]. Predictors included TA, TS, and SWin. The agreement metric and complete results are provided in Supplementary Materials.

The regression model developed from the combined tower data was applied to ERA5-Land reanalysis data to estimate monthly NEE across boreal Canadian peatlands. Predicted NEE was subsequently averaged over time to obtain a single mean growing-season value per pixel.

### Conservation index construction

To align rasters, all datasets were projected to a common coordinate system (WGS 1984), clipped to the boreal peatland extent, and resampled to the GLWD v2 grid to enable pixel-wise calculations in ArcGIS. This does not add spatial detail to predicted NEE, because all GLWD pixels within a given ERA5 cell share the same predicted NEE rate per m². Accordingly, predicted NEE and derived indices should be interpreted at the ERA5-Land resolution. All final map outputs were reprojected to the Canada Lambert Conformal Conic projection for presentation.

The Predicted Carbon Sink Strength (PCSS) represents the predicted strength of $CO_2$ sinks (NEE<0) across boreal peatlands, scaled by the GLWD v2 peatland fractional coverage. To compute PCSS, we multiplied predicted NEE by the peatland fraction within each pixel. To prioritize strong $CO_2$ sinks, we set positive NEE values to zero before calculating PCSS, thereby excluding carbon-source pixels. The resulting layer was then multiplied by -1 to make sink strength positive, and min–max normalization was applied (Equation 2):

**Equation 2:**

$$PCSS_{norm} = \frac{PCSS - \min(PCSS)}{\max(PCSS) - \min(PCSS)} \cdot 100$$

The Inverted Human Footprint (IHFP) visualizes human pressure across peatlands within our study area. We used the 2020 Global Terrestrial Human Footprint dataset (HFP-100)[12], which maps the intensity of human pressure at ~100 m resolution. HFP-100 is a composite index of anthropogenic pressure derived from multiple infrastructure, land-use, and population-related indicators.

To quantify human pressure for conservation prioritization, all Canadian tiles were extracted and mosaicked to produce a single seamless raster. To ensure that highly remote pixels received the greatest values, the raster was inverted. The resulting IHFP layer was then normalized using min–max scaling (Equation 3):

**Equation 3:**

$$IHFP_{norm} = \frac{IHFP - \min(IHFP)}{\max(IHFP) - \min(IHFP)} \cdot 100$$

The BPCI was produced by combining IHFP and PCSS using a weighted scheme (Equation 4). We selected weights through sensitivity testing to balance predicted sink strength with low disturbance, with slightly higher weight on IHFP to retain visibility of human-pressure gradients.

**Equation 4:**

$$BPCI = 0.6 \cdot IHFP_{norm} + 0.4 \cdot PCSS_{norm}$$

## Results

### Peatland coverage and site characteristics

Figure 1 shows the boreal peatland extent, fractional peatland cover, and the four AmeriFlux sites. Peatland cover is most contiguous in the Hudson Bay Lowlands, with additional high-cover regions in parts of the Northwest Territories; fractional cover is lower in mountainous regions and across much of eastern boreal Canada.

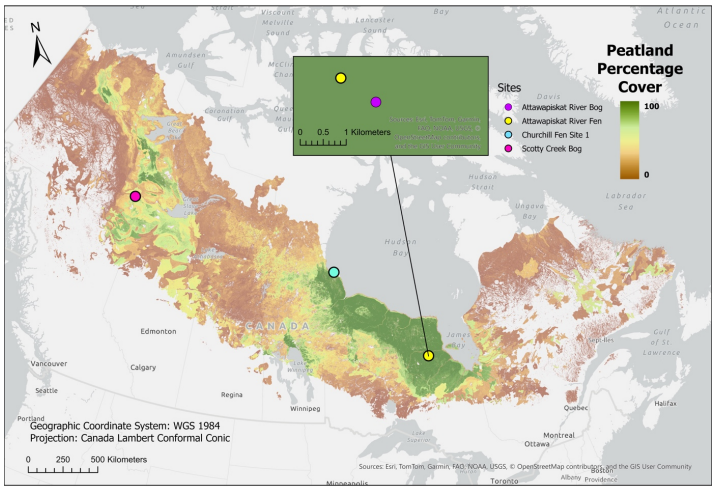



**Figure 1.** Study area, boreal peatland extent, and flux-tower sites. Map of the Canadian boreal domain showing boreal peatland fractional cover13, the boreal extent14, and four AmeriFlux eddy-covariance sites (CA-ARB, CA-ARF, CA-CF1, CA-SCB)15. The green inset enlarges the Attawapiskat River area in the Hudson Bay Lowlands to distinguish the two closely spaced sites, CA-ARB and CA-ARF; the inset scale bar indicates local distance. Map shown in the Canada Lambert Conformal Conic projection.

### Exploratory analysis of climatic predictors

Pairwise correlations among candidate predictors were assessed to screen for collinearity. TA and LWin were strongly correlated (ρ=0.92), exceeding the a priori threshold (|ρ|>0.8). We excluded LWin from subsequent analysis and regression modeling and retained TA as the more interpretable predictor. Multivariable collinearity among the remaining predictors was low (all VIFs<5), supporting their inclusion in candidate models (Supplementary Figure S2). VIF values for the final predictor set were 4.98 for TA, 2.99 for subsurface TS, and 3.25 for SWin.

Using pooled growing-season data, we examined correlations between NEE and the retained predictors TA, TS, P, and SWin (Figure 2). TA (Figure 2A) and TS (Figure 2C) were negatively correlated, with NEE becoming more negative as temperatures increased. The association was stronger for TA than TS (ρ=−0.79 versus ρ=−0.39). The relationship between SWin and NEE was also negative (Figure 2D, ρ=−0.59). Precipitation (log-transformed; Figure 2B) showed no evidence of a significant association with NEE.

| Name | Attawapiskat River Bog | Attawapiskat River Fen | Churchill Fen Site 1 | Scotty Creek Bog |
|---|---|---|---|---|
| Site ID | CA-ARB | CA-ARF | CA-CF1 | CA-SCB |
| Wetland Type | Bog | Fen | Fen | Bog |
| Biome Classification | Boreal | Boreal | Boreal | Boreal |
| Permafrost Influence | Discontinuous | Discontinuous | Continuous | Discontinuous |
| Forest Influence | Partially Forested | Partially Forested | Unforested | Unforested |
| Mean Monthly Precipitation [mm] | 700 | 700 | 452 | 388 |
| Mean Monthly Temperature [°C] | -1.3 | -1.3 | -6.5 | -2.8 |
| Elevation [m] | 90 | 88 | 16.5 | 280 |
| Start Year of Measurements | 2011 | 2011 | 2007 | 2014 |
| End Year of Measurements | 2015 | 2015 | 2008 | 2019 |
| Latitude | 52.70 | 52.70 | 58.67 | 61.31 |
| Longitude | 83.96 | 83.96 | 93.83 | 121.30 |

**Table 1.** Summary of the four Canadian boreal peatland AmeriFlux site characteristics and data coverage.

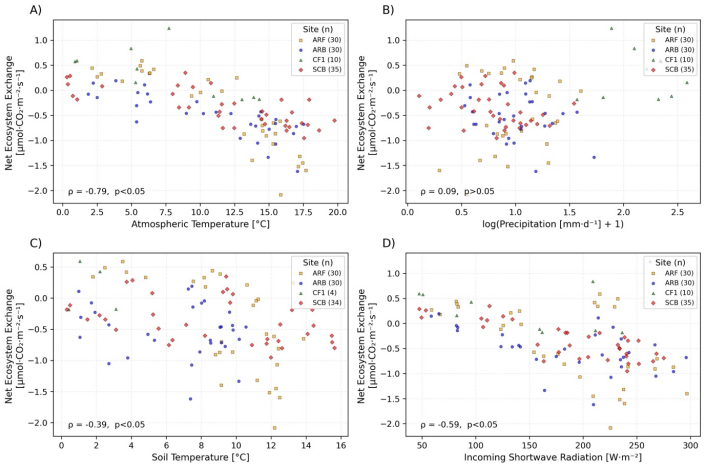



**Figure 2.** Net ecosystem exchange in relation to climatic predictors across sites during the growing season (TA>0 °C), using pooled monthly data from four AmeriFlux boreal peatland sites. Points are coloured by site. Negative NEE indicates net $CO_2$ uptake.

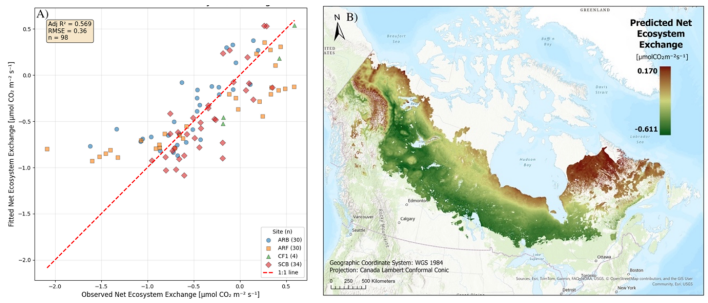



**Figure 3.** Regression model performance and regional upscaling of growing-season NEE. (A) Observed versus predicted monthly NEE from the pooled multiple linear regression (n=98; adjusted $R^2$=0.57; RMSE=0.36 μmol $CO_2$ $m^{-2}$ $s^{-1}$) using data from four boreal peatland AmeriFlux sites (CA-ARB, CA-ARF, CA-CF1, CA-SCB). (B) Spatially upscaled mean NEE (μmol $CO_2$ $m^{-2}s^{-1}$) predicted by applying the fitted model to ERA5-Land predictors across boreal Canadian peatlands. Negative values indicate net $CO_2$ uptake.

The AmeriFlux observations were used to fit a pooled multiple linear regression model to quantify the effects of the selected climatic predictors on growing-season NEE. The final regression used for spatial upscaling with ERA5-Land reanalysis data retained TA, TS, and SWin.

### The fitted NEE regression model

**Equation 5:**

$$NEE_{pred} = 0.7160 - 0.0536 \cdot TA - 0.0023 \cdot SWin - 0.0124 \cdot TS$$

When fit to a pooled growing-season dataset of 98 monthly observations, the model explained 57% of the variability in NEE (adjusted $R^2$=0.57) and had an RMSE of 0.36 μmol $CO_2$ $m^{-2}s^{-1}$ (Figure 3A). Among the variables included, TA (β=−0.0536, p=0.001) and SWin (β=−0.0023, p=0.025) were significant predictors, whereas TS was not (β=−0.0124, p=0.423).

The predicted multi-year mean growing-season NEE ranged from −0.611 to 0.170 μmol $CO_2$ $m^{-2}s^{-1}$ across boreal peatlands (mean=−0.305; SD=0.120). The strongest sinks occur in the interior, particularly in the Northwest Territories, with sink strength weakening toward the northern boreal margin and mountainous/northeastern regions (Figure 3B).

We next assessed agreement between tower-observed and ERA5-Land predictors using percent differences (Figure 4). SWin showed the strongest agreement, whereas TA and TS exhibited modest, seasonally varying biases that were most pronounced in shoulder months. Full agreement results are shown in Supplementary Materials.

### Boreal Peatland Conservation Index and component layers

A map of normalized PCSS, with higher values indicating higher predicted sink strength, is depicted in Figure 5A. PCSS scales predicted NEE by peatland fractional cover and exhibits pronounced regional variation (mean=18.84; *SD*=17.89). High PCSS values are concentrated in the central Hudson Bay Lowlands and parts of northern Manitoba and the Northwest Territories, particularly between Great Bear Lake and Great Slave Lake. By contrast, eastern regions, including northern Quebec and Labrador and mountainous areas, show lower PCSS values, with many pixels near zero. Figure 5B maps IHFP, with higher values indicating lower human pressure. IHFP is broadly high across peatland regions, reflecting generally low human pressure overall (mean=99.30; *SD*=2.59), with localized low-IHFP corridors and nodes concentrated along the southern boreal margin. These features align with major transportation routes, settlement clusters, and resource-extraction areas, where development pressure is pronounced.

Finally, Figure 5C shows BPCI values across Canadian boreal peatlands, with higher values indicating higher PCSS and lower human footprint. The largest contiguous areas of high BPCI values occur across the southern Hudson Bay Lowlands. High BPCI values also appear in the Northwest Territories, between Great Bear Lake and Great Slave Lake, and north of Lake Winnipeg. Low BPCI occurs along the southern margin, near corridors/settlements, and across much of eastern boreal Canada.

## Discussion

This study develops a climate-driven regression model that upscales growing-season peatland NEE and combines an IHFP layer with predicted $CO_2$ sink potential to produce a conservation index. By linking tower-derived relationships to ERA5-Land predictors, the workflow produces interpretable first-pass maps to support conservation and land-

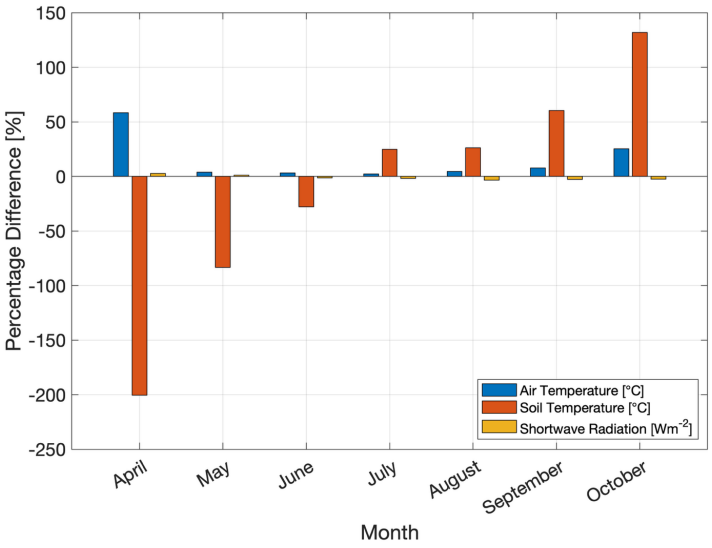



**Figure 4.** Agreement between tower-observed and ERA5-Land predictors during the growing season (TA>0 °C). Percent difference between ERA5-Land and tower-observed monthly growing-season predictors (TA, TS, SWin) at the four study sites. Positive values indicate ERA5-Land exceeds tower observations; negative values indicate underestimation.

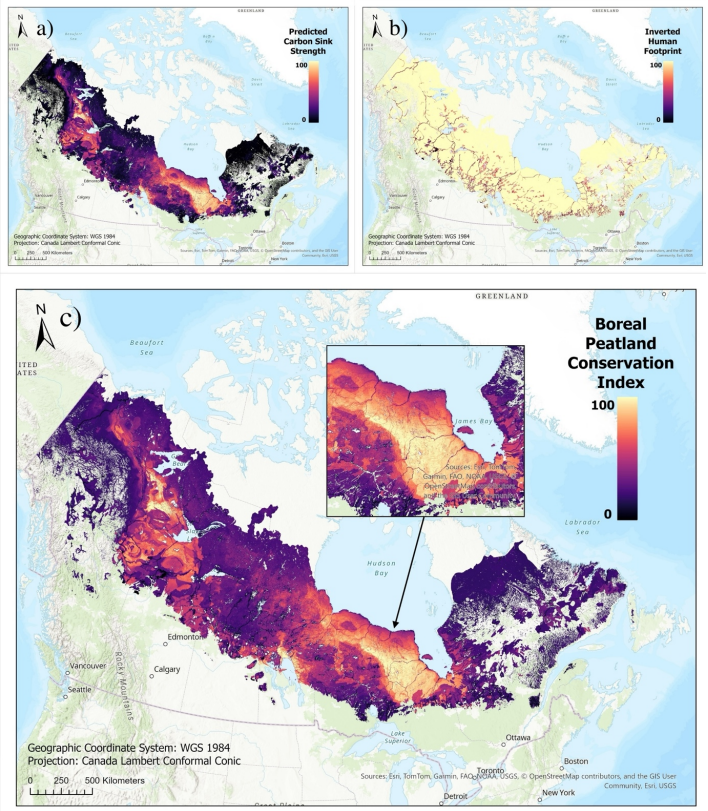



**Figure 5.** Boreal Peatland Conservation Index and component layers. (A) Predicted carbon sink strength (PCSS), derived from predicted NEE after setting source pixels (NEE>0) to zero, scaling by peatland fractional cover, and applying min–max normalization. (B) Inverted human footprint (IHFP), derived by inverting human footprint values and applying min–max normalization so that higher values indicate lower human pressure. (C) Boreal Peatland Conservation Index, calculated as a weighted combination of normalized IHFP and normalized PCSS.

use planning. This empirical approach offers a transparent complement to process-based national peatland frameworks[11]. The model explained 57% of monthly growing-season NEE variance, supporting its use for first-pass prioritization. While these maps are intended to help target consultation and research, conservation decisions should incorporate field validation, land tenure, Indigenous governance and knowledge, local governments, and current and proposed development.

## Climatic controls on growing-season NEE

Across observations, NEE was negatively correlated with TA and SWin, with a weaker but significant negative association with TS, reflecting stronger net $CO_2$ uptake during warmer, brighter months. Similarly, SWin captures an energy-availability signal in line with radiation controls on photosynthesis. Precipitation was not significantly associated with NEE and was therefore not retained for multivariable modeling.

In the multivariable model, TA and SWin best explained growing-season NEE and remained significant predictors, indicating that temperature and radiation capture much of the predictable variability, as in other regional NEE upscaling studies[7,8]. TS, while negatively correlated with NEE in pairwise analyses, was not significant in the multivariable model, likely because subsurface TS is less directly linked to near-surface fluxes and shares seasonality with other predictors. However, excluding TS reduced adjusted $R^2$, suggesting a modest complementary contribution and supporting its inclusion in the model.

Although TA and TS remained moderately correlated because of shared seasonal structure, VIF values indicated that multivariable redundancy in the retained predictor set remained within acceptable limits.

Upscaling assumes these climate-NEE relationships transfer across boreal peatlands. Figure 2 supports a first-order pooled model, but predictions are most reliable within tower-observed predictor ranges and become less certain under extrapolation. Given that adjacent CA-ARF/CA-ARB towers contribute a large share of the observations (61%), the fitted relationships may be weighted toward conditions at these sites, underscoring the value of expanding flux coverage across the boreal region.

Calibrated on four AmeriFlux peatland sites spanning bog and fen systems across discontinuous to continuous permafrost settings, the model represents only the conditions sampled by these towers. In particular, the calibration dataset does not explicitly capture gradients in nutrient status, vegetation structure (open, treed, or forested systems), or the diversity of permafrost–hydrologic regimes present across the boreal region. Process-based national modeling indicates that peatland $CO_2$ exchange varies across these ecosystem classes, with typically stronger uptake associated with forested peatlands and nutrient-rich fen systems[11]. Regional climatic gradients may also influence the transferability of climate–NEE relationships. For example, differences in moisture balance between western and eastern boreal regions can affect peatland productivity and respiration dynamics[11]. Consequently, the pooled regression approach captures first-order climatic controls on NEE but may not fully represent ecological and regional variability across all Canadian boreal peatlands.

## Reanalysis agreement and uncertainty

Tower-reanalysis comparisons show good agreement for TA and SWin during core growing-season months. Agreement weakens for TA in shoulder months, and TS shows larger, more seasonally variable percent differences. Because tower temperatures can approach 0 °C in shoulder months, small denominators can inflate percent differences, so temperature-based percent differences should be interpreted cautiously. Given the coarse ERA5-Land grid (~9 km) relative to tower footprints and sub-pixel heterogeneity

in peatlands and climate, our maps are best interpreted for regional screening rather than pixel-scale flux estimates; additional detail is provided in Supplementary Materials.

### Spatial patterns in sink potential

With these uncertainties in mind, the upscaled maps show clear regional structure in sink strength. Predicted NEE and PCSS indicate strong sink potential in the southern Hudson Bay Lowlands and parts of northern Manitoba and the Northwest Territories, weakening toward mountainous or low-cover regions and the northern margin, likely reflecting lower temperatures and reduced energy availability at higher latitudes. Weighting by peatland fractional cover highlights high-cover regions but introduces uncertainty associated with wetland classification[13]. In addition, data alignment and resampling choices can smooth local extremes, and our sink-focused PCSS approach sets positive NEE to zero, omitting information on predicted $CO_2$ sources.

Broad spatial patterns in our NEE predictions are consistent with national peatland simulations from the Canadian Model for Peatlands (CaMP)[11]. Although CaMP is process-based, both analyses highlight the Hudson Bay Lowlands as major peatland sink regions and indicate comparatively weaker sink signals across eastern boreal Canada and mountainous areas[11]. We interpret this agreement as convergence in regional patterns that supports the use of our maps for preliminary conservation screening.

### Conservation prioritization and feasibility

The BPCI combines PCSS with human footprint to highlight areas with strong predicted sinks and low disturbance. These priority areas cluster in the southern Hudson Bay Lowlands, with additional hotspots in the Northwest Territories and northern Manitoba. The BPCI can help flag areas where protecting strong carbon sinks may warrant assessment of conservation opportunities and potential project impacts. The high-scoring area in the southern Hudson Bay Lowlands appears largely spatially separated from the Ring of Fire mines and, apart from the Victor Diamond Mine, occurs in highly remote areas, which may increase the feasibility of proactive protection[22]. However, because this region overlaps with Indigenous communities[22], BPCI should be treated as a screening input for partnership-led planning and validation rather than a standalone prescription. By contrast, the northern Manitoba hotspot overlaps more strongly with existing human pressure, implying greater conflict and lower near-term feasibility than similarly strong sinks in lower-pressure areas.

### Limitations and future directions

Key considerations include sparse flux coverage and limited representation of hydrologic controls that strongly regulate peatland respiration and productivity[23]. $CO_2$ emissions can continue during the non-growing season and may offset part of the growing-season uptake[24]. Accordingly, our results represent growing-season $CO_2$ sink potential rather than the annual net carbon balance. Additionally, NEE describes ecosystem-atmosphere $CO_2$ exchange but is missing disturbance losses such as wildfire. Process-based peatland modeling indicates that wildfire can substantially negate long-term carbon sinks when moving from ecosystem productivity metrics to net biome productivity[11]. Given the importance of water-table depth for peatland $CO_2$ exchange and its coupling to $CH_4$ dynamics[10, 23, 25], future work should test spatial hydrology proxies such as drought code indices or remotely sensed soil moisture products[26]. As a multiple linear regression, the approach may under-represent non-linear or threshold responses in climate-NEE relationships. AmeriFlux gap-filling improves completeness but can introduce additional uncertainty, especially where data gaps are long. Subsurface soil-temperature harmonization may also introduce uncertainty because tower sensor depths vary across sites and do not perfectly match ERA5-Land layers. Monthly aggregation and time averaging supports regional comparability but can smooth short-lived climatic extremes that affect $CO_2$ exchange. Because agreement was assessed using percent differences, which can be inflated near 0 °C, reporting absolute bias alongside percent differences would further clarify agreement patterns. Bias-correcting ERA5 meteorological inputs to better match tower-observed climate variables, along with sensitivity analyses to quantify how ERA5 biases propagate into NEE predictions, would help bound uncertainty in flux estimates. Additional flux sites, longer time series, and annual carbon-balance estimates would increase representativeness and confidence in extrapolation. Because our calibration period spans 2007–2019, future work should incorporate newer observations as climate conditions evolve; extending the framework to include climate projections could help evaluate the persistence of current sink hotspots under a warming climate. Model evaluation would also benefit from leave-one-site-out cross-validation as additional sites become available.

Regarding disturbance and implementation feasibility, HFP-100 may under-represent fine-scale or rapidly changing disturbances, including emerging resource extraction sites. Because $CO_2$-only prioritization can misrepresent net climate benefits in methane-rich systems[2], future work should integrate $CH_4$ alongside $CO_2$-based NEE. Alternative weightings and additional data layers could be incorporated to reflect different policy goals while maintaining the workflow. Adding wildfire risk, permafrost vulnerability, and proximity to planned development could better target lasting climate benefits.

## Conclusion

Taken together, this study offers a framework for identifying boreal peatland regions where protecting $CO_2$ sinks could yield substantial climate benefits with minimal development conflict. The generated maps indicate strong $CO_2$ uptake in parts of the Northwest Territories and, after accounting for peatland fractional cover, highlight potential high-priority areas within extensive peatland complexes, including the southern Hudson Bay Lowlands. By combining sink strength with low human pressure, the Conservation Index offers a practical screening tool to support targeted field validation, consultation, and further modeling. Future extensions that incorporate $CH_4$ fluxes, hydrologic controls, permafrost dynamics, and more recent data will improve estimates of net climate benefits as climate and land-use pressures evolve.

## Acknowledgements

This research was made possible through the guidance, mentorship, and encouragement of Dr. Geneviève Ali and Dr. Peter Douglas.

## Supplementary Materials

Supplementary material referenced in the text of this article may be found online at https://doi.org/10.26443/msurj.v21i1.412.

## References

1. Goodday, V., Harris, L. & Tanguay, L. *Assessing Peatland Law and Policy Across Canada: Is Canada Fulfilling Its Critical Stewardship Role?* (Wildlife Conservation Society Canada, 2024).

2. Charman, D. J. in *Encyclopedia of Inland Waters* (eds Likens, G. E.) 541–548 (Academic Press, 2009).

3. Harris, L. I. et al. The essential carbon service provided by northern peatlands. *Frontiers in Ecology and the Environment* **20,** 222–230 (2022). 10.1002/fee.2437

4. Tarnocai, C. The effect of climate change on carbon in Canadian peatlands. *Global and Planetary Change* **53,** 222–232 (2006). 10.1016/j.gloplacha.2006.03.012

5. Wildlife Conservation Society Canada. *Canadian policy failures putting globally important carbon-rich peatland ecosystems at risk, says new report* 2024. https://wcscanada.org/newsroom/news/canadian-policy-failures-putting-globally-important-carbon-rich-peatland-ecosystems-at-risk-says-new-report/.

6. Reichle, D. E. in *The Global Carbon Cycle and Climate Change* (eds ) 119–156 (Elsevier, 2020).

7. Lu, Y., Huang, Y., Jia, Q. & Xie, Y. Capturing the net ecosystem CO2 exchange dynamics of tidal wetlands with high spatiotemporal resolution by integrating process-based and machine learning estimations. *Agricultural and Forest Meteorology* **352,** 110045 (2024). 10.1016/j.agrformet.2024.110045

8. Fei, X. et al. Eddy covariance and biometric measurements show that a savanna ecosystem in Southwest China is a carbon sink. *Scientific Reports* **7,** 41025 (2017). 10.1038/srep41025

9. Wood, D. A. Net ecosystem exchange comparative analysis of the relative influence of recorded variables in well monitored ecosystems. *Ecological Complexity* **50,** 100998 (2022). 10.1016/j.ecocom.2022.100998

10. Strachan, I. B., Pelletier, L. & Bonneville, M.-C. Inter-annual variability in water table depth controls net ecosystem carbon dioxide exchange in a boreal bog. *Biogeochemistry* **127,** 99–111 (2016). 10.1007/s10533-015-0170-8

11. Bona, K. A. et al. Using the Canadian Model for Peatlands (CaMP) to examine greenhouse gas emissions and carbon sink strength in Canada's boreal and temperate peatlands. *Ecological Modelling* **490,** 110633 (2024). 10.1016/j.ecolmodel.2024.110633

12. Mazzariello, J. & Gassert, F. *Global 100m Terrestrial Human Footprint (HFP-100)* (Dryad, 2023).

13. Lehner, B. et al. Mapping the world's inland surface waters: an upgrade to the Global Lakes and Wetlands Database (GLWD v2). *Earth System Science Data* **17,** 2277–2329 (2025). 10.5194/essd-17-2277-2025

14. Brandt, J. P. The extent of the North American boreal zone. *Environmental Reviews* **17,** 101–161 (2009). 10.1139/A09-004

15. Chu, H. et al. AmeriFlux BASE data pipeline to support network growth and data sharing. *Scientific Data* **10,** 614 (2023). 10.1038/s41597-023-02531-2

16. Reichstein, M. et al. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology* **11,** 1424–1439 (2005). 10.1111/j.1365-2486.2005.001002.x

17. Pastorello, G. et al. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data* **7,** 225 (2020). 10.1038/s41597-020-0534-3

18. Lasslop, G. et al. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. *Global Change Biology* **16,** 187–208 (2010). 10.1111/j.1365-2486.2009.02041.x

19. Hersbach, H. et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146,** 1999–2049 (2020). 10.1002/qj.3803

20. Li, M. et al. Key Environmental and Ecological Variables of Wetland CH4 and CO2 Fluxes Change With Warming. *Earth's Future* **13,** e2024EF005751 (2025). 10.1029/2024EF005751

21. Petzoldt, T. *Data Analysis with R: Selected Topics and Examples* 2020. https://wwwpub.zih.tu-dresden.de/~petzoldt/elements_en.pdf.

22. Gamble, J. *What's at stake in Ontario's Ring of Fire* 2017. https://canadiangeographic.ca/articles/whats-at-stake-in-ontarios-ring-of-fire/.

23. Lafleur, P. M., Roulet, N. T., Bubier, J. L., Frolking, S. & Moore, T. R. Interannual variability in the peatland-atmosphere carbon dioxide exchange at an ombrotrophic bog. *Global Biogeochemical Cycles* **17,** 1036 (2003). 10.1029/2002GB001983

24. Rafat, A. et al. Non-growing season carbon emissions in a northern peatland are projected to increase under global warming. *Communications Earth and Environment* **2,** 1–10 (2021). 10.1038/s43247-021-00184-w

25. Strack, M. & Waddington, J. M. Response of peatland carbon dioxide and methane fluxes to a water table drawdown experiment. *Global Biogeochemical Cycles* **21,** GB1007 (2007). 10.1029/2006GB002715

26. Copernicus Climate Change Service. *Fire danger indices historical data from the Copernicus Emergency Management Service* 2019.

Submitted: 02/12/2024
Accepted: 02/17/2025
Published: 04/10/2026

Research Article

[1]Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Keywords

MAMLD1, positive selection, pleiotropy, primate evolution, hypospadias

Email Correspondence

yunhua.ren@mail.utoronto.ca

https://doi.org/10.26443/msurj.v21i1.227



Ren, Yunhua[1]

# Positive Selection on the Primate MAMLD1 Gene Suggests Adaptive Evolution in a Pleiotropic Regulatory Factor

## Abstract

Gene regulatory factors (GRFs) are essential regulators of human phenotypic diversity. Among these, MAMLD1 acts as a transcriptional co-activator critical for male genital development. Mutations in *MAMLD1* are clinically associated with X-linked hypospadias type 2, a condition that alters the placement of the male urethral opening and potentially impacts reproductive fitness. In evolutionary biology, natural selection is often detected by comparing the rates of non-synonymous (*dN*) and synonymous (*dS*) substitutions ($\omega = dN/dS$). Positive selection ($\omega > 1$) represents the rapid fixation of beneficial mutations that enhance an organism's fitness, distinguishing it from purifying selection ($\omega < 1$), which removes deleterious mutations, and neutral evolution ($\omega = 1$), which involves random genetic drift. Building upon broader genomic surveys, this study provides a high-resolution analysis of *MAMLD1* using the PAML v4.9 suite across 20 primate species. By employing site models to assess selection across the primate phylogeny, we identified significant signals of positive selection. Findings suggest that *MAMLD1* has undergone adaptive evolution, likely driven by sexual selection and its role in reproductive isolation. This research underscores the gene's multifaceted role in shaping human-specific developmental phenotypes and evolutionary strategies.

## Introduction

Sequence differences in gene coding regions and variance in gene regulatory systems may both contribute to phenotypic differences between humans and other species [1]. Variations in gene regulatory mechanisms can arise from changes in the DNA sequence of a regulatory region, which may impact gene expression, or changes in sequences of so-called gene regulatory factors (GRFs), which may affect the expression of their target genes [2]. GRFs play a variety of roles in gene regulation, including regulating the timing and tissue-specificity of a gene's expression. In this study, I focused on MAMLD1, a transcriptional co-activator expressed during male urogenital development and implicated in disorders of sex development. I selected MAMLD1 because (i) it is a pleiotropic regulatory factor with multiple downstream targets, and (ii) prior comparative genomic work has suggested that some GRFs show signatures of adaptive evolution in the human lineage [3]. I therefore tested the specific hypothesis that a subset of codons in MAMLD1 experienced episodic positive selection across the primate phylogeny, potentially reflecting selection on reproductive or developmental phenotypes. Hypospadias type 2, a disorder of sex development associated with MAMLD1 mutations, is characterized by abnormal placement of the male urethral opening on the underside of the penis [4]. This condition can lead to challenges in the normal function of male genitalia, including issues with urination and, in severe cases, reproductive difficulties [5]. Given its pleiotropic properties, regulatory control over multiple genes, and essential role in normal male genital development, I hypothesized that MAMLD1 is subject to evolutionary constraint, particularly within its functional domains. These constraints are likely to preserve the integrity of crucial sequences, maintaining conserved functions across species due to the gene's essential roles in development and physiological processes [1]. It has also been proposed that, in contrast to its target genes, gene regulatory systems evolve under less selective restrictions than their target genes [1].

The significant differences between humans and great apes, as well as between great apes and other primates provide a framework for investigating the genetic mechanisms underlying interspecific variation [6]. Recent studies have shown an average genetic difference of 1.2% between humans and chimpanzees, with non-coding regions showing slightly higher genetic differences than coding regions [7]. The difference is between 3-4% when non-alignable regions of the genome, such as insertions, deletions, and rearrangements, are taken into consideration. The average genetic difference between humans and the rhesus macaque, a primate species that is more evolutionarily distant from us, was determined to be significantly greater at 6.46%, or up to 9.24% when minor insertions and deletions are taken into account [8]. Some of the sequence changes might result from neutral evolution, which involves random genetic mutations that do not confer any selective advantage or disadvantage. Other changes might be caused by ongoing adaptive interactions between genomes and the environment, leading to positive selection [9]. Under positive selection, genetic variants that increase an organism's fitness or reproductive success increase in frequency over time. Depending on the dating technique, the human lineage diverged from its close relatives, chimpanzees, and bonobos, about 5.5–11.5 million years ago (Ma) [10]. The resulting human phenotype has historically been attributed to continual adaptations to niches and local environments [11]. The discovery of genes that have evolved by positive selection can shed light on how animals adapt to their environment and provide answers to some crucial biological problems, such as how a particular phenotype developed [8]. Across primates, many coding genes are dominated by purifying selection, although the identity and strength of positive selection can differ among lineages depending on ecological, demographic, and reproductive factors. Purifying selection is the evolutionary process that eliminates deleterious alleles from a population, thereby preserving the fitness of the species by preventing harmful genetic variations from becoming widespread. It has long been hypothesized that identifying genes that have undergone positive selection along the human lineage could contribute to uncover biologically significant genetic changes that uniquely identify humans [12]. As a result, researchers have used various methods to search the human genome for signs of positive selection in particular genes. Notably, Jovanovic et al. (2021) re-

ported evidence of positive selection in MAMLD1 using expanded genomic sampling and complementary methods, identifying candidate codons such as 726 and 728. Following the findings of Jovanovic et al. (2021), the goal of the present study is to independently assess how animal model choice affects inference of the previously identified positively selected sites using a streamlined PAML analysis**.** Shared environmental challenges, known as selective pressures, can lead to similar evolutionary responses across a group of species [13], resulting in uniform selection patterns, or traits that are consistently favored by natural selection across these species. This phenomenon suggests that certain traits may emerge not by coincidence but as adaptive responses to similar environmental demands. However, positively selected mutations are seldom found at polymorphic sites [3]. Through the process of positive selection, beneficial mutations are typically fixed rapidly within a population. This rapid fixation means that advantageous genetic variations tend to become common or even universal among individuals of a species, as they confer an increased chance of survival and reproduction. The exclusion of probable false positive candidates discovered at the species level is thus made possible by examining the polymorphism of positively selected codons at the population level. The main objective of this paper is to examine the MAMLD1 gene as a potential candidate that may exhibit signs of positive selection, with a focus on how selective pressures on MAMLD1 are distributed across primate species, supporting the gene's functional importance in human evolution. The ability to detect positive selection within a gene is significantly influenced by the number of available sequences for analysis. A larger dataset of sequences, including those from humans and their close relatives, enhances the statistical power of evolutionary models to identify subtle changes indicative of positive selection. Consequently, this study utilizes 20 sequences, encompassing both human and closely related species, to ensure a robust analysis capable of detecting signs of positive selection on the MAMLD1 gene.

Importantly, MAMLD1 is located on the X chromosome, which can experience evolutionary dynamics that differ from autosomes due to hemizygosity in males, sex-biased effective population size, and potentially greater exposure of recessive alleles to selection. In addition, human evolutionary history includes demographic bottlenecks and genetic drift, which can reduce genetic diversity and sometimes mimic or obscure signals of selection in sequence-based tests. Because codon models infer selection from patterns of fixed substitutions along branches, demographic history is not modeled explicitly here; therefore, lineage-level evidence should be interpreted cautiously and ideally complemented by population-level tests of fixation and by comparison to control loci.

**Table 1.** Summary of sequence data used. Species name, corresponding MAMLD1 accession number on NCBI and abbreviation. The first 10 sequences are the original dataset provided by Prof. Irwin; the last 10 sequences are displayed in search order.

| Species name | Accession number | Abbreviation |
|---|---|---|
| Homo sapiens | ENST00000370401 | Human |
| Pan troglodytes | ENSPTRT00000107245 | T_glodytes |
| Nomascus leucogenys | ENSNLET00000017270 | L_cogenys |
| Macaca mulatta | ENSMMUT00000074913 | Mulatta |
| Microcebus murinus | ENSMICT00000017610 | Murinus |
| Otolemur garnettii | ENSOGAT00000013472 | Garnettii |
| Cebus imitator | ENSCCAT00000018615 | Imitator |
| Rhinopithecus roxellana | ENSRROT00000058882 | Roxellana |
| Aotus nancymaae | ENSANAT00000034312 | Nancymaae |
| Pongo abelii | ENSPPYT00000024273 | Abelii |
| Pan paniscus | XM_008968068 | Paniscus |
| Gorilla gorilla gorilla | XM_004065005 | Gorilla |
| Papio anubis | XM_031661028 | Anubis |
| Trachypithecus francoisi | XM_045384134 | Francoisi |
| Cercocebus atys | XM_012081351 | Atys_dna |
| Macaca thibetana thibetana | XM_050775309 | Thibetana |
| Hylobates moloch | XM_032757793 | Moloch |
| Macaca nemestrina | XM_011765830 | N_estrina |
| Macaca fascicularis | XM_045384134 | F_icular |
| Rhinopithecus bieti | XM_017865534 | Bieti |

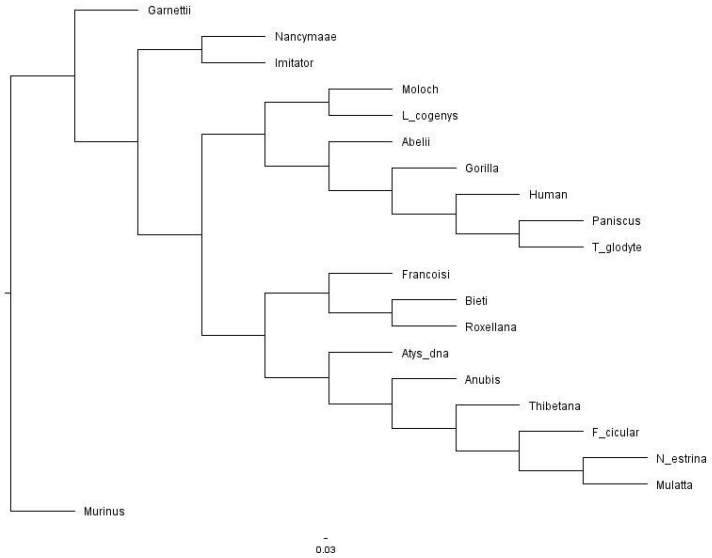



**Figure 1.** . Primate species tree used for codon-model analyses. Species are labeled using common names (human, chimpanzee, bonobo, gorilla, orangutan, baboon, rhesus macaque, etc.) to improve readability. *Microcebus murinus* is used to root the primate tree

## Materials and Methods

### Sequence Dataset and phylogenetic analysis

A total of 20 sequences were used for this analysis, one human sequence and 19 ortholog sequences from closely related primates. First ten sequences in Table 1 were downloaded from Ensembl and the last ten sequences were from the NBCI nBLAT website. Alignment was performed using MUSCLE (codon-aware) in MEGA, and the phylogeny was inferred using PhyML; FigTree was used only for tree visualization. The TN93 + G model of evolution was chosen as it was the best fit model according to AICc values. This analysis uses one curated reference coding sequence per species (i.e., an orthologous representative) to infer long-term substitution patterns along branches of the species tree. In codon-based phylogenetic models, branch lengths summarize substitutions accumulated since divergence and therefore reflect lineage-level evolutionary history rather than within-species polymorphism. Consequently, these models can detect recurrent selection operating over millions of years, but they cannot determine whether candidate sites are fixed or polymorphic within modern populations. We therefore interpret lineage-level signals using log-likelihood ratio tests (LRTs) based on ω as evidence of historical selection pressures and treat population-level verification as an important limitation and future direction. In general, the above displayed tree (Figure 1) has a similar topology to a recently published species tree [14].

Candidate sequences were obtained as annotated one-to-one orthologs from Ensembl when available. Additional primate sequences were retrieved from NCBI and verified by reciprocal best-hit BLAST against the human MAMLD1 coding sequence to reduce the risk of paralog inclusion.

### Site Model test for Positive Selection

Estimating the ratio of non-synonymous to synonymous substitutions (ω) between the coding regions of homologs is a common method for measuring the selective pressure acting on protein-coding genes. The empirical p-values were calculated under the assumption that the log-ratio tests (LRT) had a $\chi 2$ distribution. The Bayes Empirical Bayes (BEB) approach, implemented in PAML v4.9, was used to estimate the posterior probabilities that individual codon sites belong to site classes with $dN/dS$ (ω) > 1. This method identifies codons with strong statistical support for positive

**Table 2.** Summary table of PAML (Phylogenetic Analysis by Maximum Likelihood) analysis, showcasing the results of running several evolutionary models on the MAMLD1 gene sequences.

| Model | np | lnL | k | | | | | Null | LRT | df | P |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M0 | 39 | -3678.103535 | 1.15655 | | | | | n/a | | | |
| M1a | 40 | -35016.051963 | 0.8736 | **p:** | 0.62208 | 0.37792 | | M0 | 3564.103 | 1 | 0.0000 |
| | | | | **w:** | 0.05164 | 1.00000 | | | | | |
| M2a | 42 | -33345.590264 | 1.16499 | **p:** | 0.16244 | 0.71091 | 0.12665 | M1a | 3340.923 | 2 | 0.0000 |
| | | | | **w:** | 0.19078 | 1.00000 | 95.07202 | | | | |
| M2a_rel | 42 | -33345.590623 | 1.16499 | **p:** | 0.16245 | 0.71090 | 0.12665 | M1a | 3340.923 | 2 | 0.0000 |
| | | | | **w:** | 0.19078 | 1.00000 | 95.07109 | | | | |
| M3 | 43 | -33345.355812 | 1.16018 | **p:** | 0.15537 | 0.71707 | 0.12756 | M0 | 6905.495 | 4 | 0.0000 |
| | | | | **w:** | 0.17866 | 0.95563 | 91.89691 | | | | |
| M7 | 40 | -35127.479377 | 0.85969 | **p:** | p=0.31920 | q=0.39792 | | n/a | | | |
| M8 | 42 | -33331.356206 | 1.13260 | **p:** | p=0.85848 | q=0.34956 | | M7 | 3592.246 | 2 | 0.0000 |
| | | | | **w:** | 79.05894 | | | | | | |

selection across the primate phylogeny. Subsequent investigations should concentrate on codons identified as being under positive selection.

M0 (one-ratio): assumes a single rate of evolution across all branches. M1a (nearly neutral): contrasting neutral versus selection sites. M2a (positive selection): allows for sites under positive selection. M2a_rel: a refined version of M2a offering more detailed insights into selection. M3 (discrete): permits variable selection across sites. M7 (beta): modeling variation under a beta distribution without positive selection. M8 (beta&ω): introducing positive selection on top of the beta distribution. Each model plays a crucial role in identifying different aspects of evolutionary pressures acting on the gene, with particular emphasis on pinpointing specific codons where positive selection may have occurred, offering insights into the adaptive evolution of the MAMLD1 gene in humans and their close relatives.

## Results

### Positive Selection Across the Primate MAMLD1 Gene

Testing for positive selection pressure was detected using codon models of evolution implemented in PAML (Yang 2007). These sites are critical positions within the gene's coding region where amino acid substitutions occur. The ω (omega) parameters, which represent the ratio of non-synonymous (*dN*) to synonymous (*dS*) substitution rates, vary among these sites. A ratio greater than 1 (ω > 1) at any site suggests that the site is under positive selection, indicating that amino acid changes at these positions are favored by natural selection because they likely confer some adaptive advantage. Our analysis revealed that under models M2a, M3, M7, and M8, several sites across the MAMLD1 gene exhibit ω parameters indicative of positive selection pressure, suggesting these coding sites have undergone adaptive changes that could be functionally relevant to organismal fitness. These adaptive changes in the protein indicate positive selection.

In Table 2, column 'p' represents the estimated proportion of sites belonging to a specific ω category, while column 'w' represents the estimated ω (*dN/dS*) value for that category. The p-value (Column M) represents the statistical significance of the Likelihood Ratio Test (LRT). Specifically, model M8 identified a class of sites with a highly elevated ω = 79.06, and model M2a identified a class of sites with ω = 95.07, indicating strong positive selection at these positions.

## Discussion

The codon-model analyses are consistent with episodic positive selection acting on a subset of MAMLD1 codons. These methods primarily detect recurrent selection that leaves multiple fixed nonsynonymous substitutions. They may miss adaptation driven by a single major change, and they cannot on their own distinguish fixation from standing polymorphism. Biologically, MAMLD1 is implicated in male urogenital development, so one plausible selective context is sexual selection or other reproductive selection pressures acting on fertility-related phenotypes. However, linking sequence evolution in a single locus to speciation requires additional evidence such as demonstrated reproductive barriers, functional effects of the substituted residues, or consistent divergence across populations. Thus we treat speciation as a hypothesis that invites further research in this area.

Although MAMLD1 has been identified as a GRF under positive selection, no interactions between MAMLD1 and other GRFs have been reported to our knowledge. This lack of reported interactions highlights the unique evolutionary pathway of MAMLD1 and underscores the importance of further research into its specific role and mechanisms within the regulatory networks influencing human development and evolution [14]. It is possible to hypothesize that this is because the set of genes it controls are separate or functioning via different pathways, and the epistasis between them cannot be identified using the data currently available. The independenceof MAMLD1's effects on different traits and its potential role in complex phenotypes can be accounted if potential pleiotropic effect is considered. Because MAMLD1 normally regulates the expression of several genes, it can have physiological and morphological implications across scales, affecting cells, tissues or at the level of whole organisms [15]. These implications could be independently adaptive; this means that different mutations or variations within this gene can lead to separate adaptive advantages. It has been shown that minor mutations, even within a single gene, may provide a quick road to phenotypic adaption [16]. The different positively selected sites (PSSs) that were found within MAMLD1 could add to the pleiotropic effect and connect with distinct phenotypes. This has been observed before for certain other genes, where various polymorphism sites have varied trait relationships [4]. Another reasonable explanation for the lack of gene interactions is that the selective pressures on MAMLD1 were acting at a distinct time points after the split of the human lineage. For MAMLD1, this could mean that the gene has undergone several rounds of optimization, each enhancing different aspects of its regulatory functions. Early in human evolution, for instance, a selective event might have favored mutations that improved reproductive success, while much later, another event might have selected for variations that enhanced the gene's role in development or disease resistance.

The evolutionary history of MAMLD1, marked by these separate yet cumulative adaptations, underscores its potential significance in understanding the genetic foundations of human development, health, and disease.

A key limitation of the current design is the absence of comparator loci that would contextualize model behavior such as an X-linked gene with no prior expectation of positive selection as a negative control, and an autosomal gene with reported positive selection as a positive control. Includ-

ing such controls would help evaluate whether the magnitude of likelihood improvements observed for MAMLD1 is exceptional or typical under the same alignment and model settings. In addition, outgroup choice can influence branch length estimates; here I used *Microcebus murinus* as an outgroup within primates to root the tree while avoiding the extreme divergence and life-history differences that would be introduced by using rodents. Future work should implement explicit control-gene analyses and sensitivity checks for outgroup selection.

## Acknowledgments

This research was conducted as part of an undergraduate course HMB460H1: Molecular Evolution and Genomics at University of Toronto St. George Campus. I would like to thank Prof. David M. Irwin and Prof. Belinda Chang for their guidance, support, and valuable insights.

## References

1. Anderson, J., Vilgalys, T. & Tung, J. Broadening primate genomics: new insights into the ecology and evolution of primate gene regulation. *Curr. Opin. Genet. Dev.* **62,** 16–22 (2020).
2. Siepel, A. & Arbiza, L. Cis-regulatory elements and human evolution. *Curr. Opin. Genet. Dev.* **29,** 81–89 (2014).
3. Jovanovic, V. et al. Positive selection in gene regulatory factors suggests adaptive pleiotropic changes during human evolution. *Front. Genet.* **12** (2021).
4. Chen, Y. Mutational study of the MAMLD1-gene in hypospadias. *Eur. J. Med. Genet.* **53,** 122–126 (2010).
5. Fukami, M. et al. CXorf6 is a causative gene for hypospadias. *Nat. Genet.* **38,** 1369–1371 (2006).
6. Nickel, G., Tefft, D., Goglin, K. & Adams, M. An empirical test for branch-specific positive selection. *Genetics* **179,** 2183–2193 (2008).
7. Kronenberg, Z. et al. High-resolution comparative analysis of great ape genomes. *Science* **360** (2018).
8. Su, Z. et al. Species specific exome probes reveal new insights in positively selected genes in nonhuman primates. *Sci. Rep.* **6** (2016).
9. Varki, A., Geschwind, D. & Eichler, E. Human uniqueness: genome interactions with environment, behaviour and culture. *Nat. Rev. Genet.* **9,** 749–763 (2008).
10. Jeong, C. & Rienzo, A. Adaptations to local environments in modern human populations. *Curr. Opin. Genet. Dev.* **29,** 1–8 (2014).
11. Clark, A. et al. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* **302,** 1960–1963 (2003).
12. Shultz, A. & Sackton, T. Immune genes are hotspots of shared positive selection across birds and mammals. *eLife* **8** (2019).
13. Slodkowicz, G. & Goldman, N. Integrated structural and evolutionary analysis reveals common mechanisms underlying adaptive evolution in mammals. *Proc. Natl. Acad. Sci. U.S.A.* **117,** 5977–5986 (2020).
14. Yang, Z. & Nielsen, R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol. Biol. Evol.* **19,** 908–917 (2002).
15. Linnen, C. et al. Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339,** 1312–1316 (2013).
16. Mackay, T., Stone, E. & Ayroles, J. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* **10,** 565–577 (2009).

Submitted: 01/06/2026
Accepted: 03/07/2026
Published: 04/10/2026

Research Artic;e

[1]Department of Psychology, University of Toronto, Toronto, ON, Canada

Keywords

conscientiousness, openness, intrinsic goals, extrinsic goals, well-being

Email Correspondence

georgia.roberts@mail.utoronto.ca

https://doi.org/10.26443/msurj.v21i1.385



Roberts, Georgia[1]

# Personality and Wellbeing: How Conscientiousness and Openness Influence Wellbeing Through Intrinsic and Extrinsic Goal Aspirations and Attainment

## Abstract

Well-being is a major focus of psychological research worldwide. Although environmental factors, personality traits, and goal achievement each predict well-being, they are often examined in isolation. However, it remains unclear how these factors are connected, for example, whether personality influences well-being in part through its impact on goal setting and goal attainment. The present study examines whether conscientiousness and openness relate to well-being through goal aspirations and goal attainment. Guided by Self-Determination Theory, intrinsic and extrinsic goals were tested as mediating pathways connecting conscientiousness and openness to well-being. Using data from 1,892 adults across 51 countries, multiple mediation analyses showed that both conscientiousness and openness were positively associated with well-being. Conscientiousness predicted greater intrinsic goal importance and intrinsic goal attainment, which partially mediated its relationship with well-being. Conscientiousness also negatively predicted extrinsic goal importance. Openness strongly predicted intrinsic goal importance and attainment, while negatively predicting extrinsic goals. The association between openness and well-being was fully mediated by intrinsic goal attainment. Extrinsic goal importance and attainment were not significantly related to well-being. Overall, these findings highlight intrinsic goal pursuit as a key psychological mechanism through which personality traits impact well-being.

## Introduction

Personality structures refer to stable patterns of behavior, thought, and emotion that vary across individuals and help explain consistent psychological functioning[1]. Among the most widely accepted models for understanding personality is the Big Five framework[2], which categorizes traits into neuroticism, extraversion, agreeableness, openness to experience, and conscientiousness. These traits reflect broad tendencies in how people think, feel, and act across different situations.

Conscientiousness is marked by self-discipline, persistence, impulse control, and the ability to pursue long-term goals. Individuals high in conscientiousness tend to plan ahead, delay gratification, and regulate impulses effectively[3]. This trait has been consistently linked to enhanced well-being via greater life satisfaction and lower rates of anxiety and depression[4, 5]. One explanation is that conscientious individuals engage in health-promoting behaviors, regulate negative emotions effectively, and use strategic planning to achieve goals[6, 7, 8, 9]. By shaping behavior and goal-directed effort, conscientiousness helps create conditions that support well-being.

Openness to experience is a personality trait characterized by imagination, curiosity, intellectual curiosity, and a preference for novelty[2]. Individuals high in openness tend to seek out new experience and ideas, and are more receptive to unconventional perspectives. Although openness shows less consistent associations with well-being than conscientiousness, research suggests it may influence well-being through indirect pathways. Individuals high in openness are more likely to engage in creative activities, pursue meaningful experiences, and explore personal interests, which can support purpose and life satisfaction[4].

The mechanisms linking personality traits, such as conscientiousness and openness, to well-being remain only partially understood. One promising explanation involves the types of goals individuals set, and their success in achieving them[10]. Personality influences both goal selection and the likelihood of attainment[9]. According to Self-Determination Theory (SDT), individuals pursue two broad types of goals: intrinsic goals, which directly satisfy basic psychological needs (relatedness, competence, and autonomy), and extrinsic goals, which are oriented toward external rewards like wealth, fame, or status. Intrinsic goal aspirations are consistently associated with higher well-being[11], whereas prioritizing extrinsic goals is often linked to lower well-being because they are less directly tied to psychological need fulfillment. However, attaining either type of goal can enhance well-being[12, 13]. This distinction reflects differences between aspiration and attainment: pursuing extrinsic goals alone may not satisfy psychological needs, whereas achieving goals can provide benefits such as financial stability or opportunities that indirectly support well-being. In contrast, intrinsic goals contribute to well-being both through the pursuit itself and through attainment, as they fulfill core psychological needs and provide meaningful experiences. Thus, the impact of a goal on well-being depends not only on its type, but also on whether it is successfully achieved.

The present study examines how conscientiousness and openness contribute to well-being through their associations with intrinsic and extrinsic goal aspirations and attainment. These traits were selected because, compared to more emotion-focused traits such as neuroticism and extraversion, conscientiousness and openness are closely related to how individuals set priorities and work toward goals. Although agreeableness is often examined within the framework of Self-Determination Theory, it was not included in the present study because our focus was specifically on traits most directly linked to the striving, persistence, and the pursuit of both intrinsic and extrinsic goals. Conscientiousness, because it is characterized

by planning and persistence, is expected to predict goal aspirations, goal attainment, and well-being. Openness is expected to predict intrinsic aspirations and well-being, as individuals high in openness often value personal growth, self-exploration, and meaningful relationships, which align with intrinsic goals[14]. It is hypothesized that attainment of both intrinsic and extrinsic goals will be positively associated with well-being, with intrinsic attainment showing a stronger effect. Additionally, conscientiousness is expected to relate to attainment of both goal types. Overall, the relationship between conscientiousness, openness, and well-being is expected to be partially mediated by goal aspirations and attainment.

## Methods

### Participants and Procedure

A total of 1,892 participants from 51 countries, predominantly the United States (83%), completed an online survey for financial compensation. Participants provided demographic information, including age, gender, ethnicity, and personal income (48% male, 51% female, 1% preferred not to say). Demographic variables were not analyzed in the presented study. All participants provided informed consent prior to participation and were free to withdraw at any time without penalty. Participant confidentiality and data privacy were maintained throughout the study. The survey was administered online using a paid Qualtrics panel. Qualtrics panels consist of pre-recruited participants who have agreed to complete surveys. When a sample is purchased, Qualtrics recruits respondents through its global panel partners and distributes the survey to them.

### Measures

Conscientiousness and openness were each assessed using 20 items from the Big Five Aspect Scales[15]. This 100-item measure includes 10 subscales, representing two aspects for each of the Big Five personality traits, with 10 items per aspect. For conscientiousness, the two aspects measured were Industriousness (e.g., diligence, self-discipline, goal striving), and Orderliness (e.g., organization, tidiness, preference for structure), with items such as 'I carry out my plans'. For openness, the aspects assessed were Openness (e.g., imagination, aesthetic sensitivity, curiosity for art/nature), and Intellect (e.g., intellectual engagement, abstract thinking, cognitive reflection), with items such as 'I have a rich vocabulary'. Items were combined into total scores for each trait to provide reliable overall measures, as our primary focus was on the broader trait rather than differences between the aspects. Participants responded to each item on a 7-point Likert scale ranging from 1 (Strongly Disagree) to 7 (Strongly Agree). Higher scores indicated greater levels of conscientiousness or openness. Half of the items for each trait (10 out of 20) were reverse-coded.

Intrinsic and extrinsic goal aspiration and attainment were each measured using 15 items each from the Aspiration Index[16]. Participants were presented with a series of life goals, one at a time, and asked two questions about each: (a) "How important is this goal to you?" and (b) "How much have you already attained this goal?" Responses were recorded on a 7-point Likert scale ranging from 1 (Not at all) to 7 (Very), with 4 indicating Moderately. Examples of goals included intrinsic goals such as "To help people in need," and extrinsic goals such as "To have an image that others find appealing." Higher scores indicated greater aspiration or attainment for a given type of goal.

Well-being was measured by a 12-item Scale of Positive and Negative Experiences[17], and a 5-item Satisfaction with Life Scale[18]. The Scale of Positive and Negative Experiences includes 6 items of each valence, for example, happy, joyful, sad, and afraid. Participants indicated on a 5-point Likert Scale the frequency of experiencing each feeling (1-Very Rarely or Never, 5-Very Often or Always). The Satisfaction with Life Scale lists statements for participants to indicate their degree of agreement with each statement via a 7-point Likert Scale (1-Strongly Disagree, 4-Neither Agree nor Disagree, 7-Strongly Agree). As an example, one item states "In most ways my life is close to my ideal." Higher scores indicated greater life satisfaction. No items were reverse-coded. To enable combination of the two measures into a single composite well-being index, scores were standardized (z-scored) prior to aggregation.

Several additional measures were also administered but not included in the data analyses, including demographic information, the Awareness of Narrative Identity Questionnaire[19], Basic Psychological Need Satisfaction and Frustration Scale[20], Perceived General Health[21], the Pennebaker Inventory of Limbic Languidness[22], and the 16-item International Cognitive Ability Resource[23]. This dataset was collected in collaboration with other researchers with the intention of testing multiple hypotheses; however, only the measures relevant to the present hypotheses were utilized in this study. Although basic psychological need satisfaction is considered highly related to well-being within the framework of Self-Determination Theory, it was not included in the present analyses because it was outside the scope of the specific research questions addressed in this study.

**Table 1.** Summary of Variable Means, Standard Deviations, Internal Consistency and Bivariate Correlations

| Variable | M | SD | Cronbach's α | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Wellbeing | 0 | 0.9 | 0.91* | — | | | | | | |
| 2. Conscientiousness | 4.77 | 0.82 | 0.88* | .43** | — | | | | | |
| | | | | [.40, .47] | | | | | | |
| 3. Openness | 5.06 | 0.77 | 0.85* | .18** | .37** | — | | | | |
| | | | | [.13, .22] | [.33, .41] | | | | | |
| 4. Intrinsic Goal Importance | 5.48 | 1.01 | 0.90* | .24** | .28** | .48** | — | | | |
| | | | | [.20, .28] | [.24, .32] | [.45, .52] | | | | |
| 5. Intrinsic Goal Attainment | 4.26 | 1.05 | 0.92* | .57** | .31** | .24** | .46** | — | | |
| | | | | [.54, .60] | [.27, .35] | [.19, .28] | [.42, .49] | | | |
| 6. Extrinsic Goal Importance | 3.48 | 1.25 | 0.87* | 0.04 | -.06* | -.12** | .23** | .18** | — | |
| | | | | [-.01, .08] | [-.10, -.01] | [-.16, -.07] | [.19, .27] | [.14, .22] | | |
| 7. Extrinsic Goal Attainment | 2.92 | 1.2 | 0.91* | .29** | -0.01 | -.15** | .06** | .55** | .66** | — |
| | | | | [.25, .33] | [-.05, .04] | [-.20, -.11] | [.02, .11] | [.51, .58] | [.63, .68] | |

**Note.** M = mean; SD = standard deviation; Cronbach's α = internal consistency of multi-item scales. Values in square brackets represent 95% confidence intervals. The confidence interval indicates a plausible range of population correlations that could have produced the observed sample correlation. *p < .05, **p < .01.

## Results

### Sample Characteristics

Conscientiousness and openness were the primary personality traits of interest in the present study. Overall, analysis conducted using R software indicated that participants reported moderately high levels of conscientiousness. Scores ranged from 1.05 to 7.00, with a mean of 4.77 (median = 4.90). Participants also reported relatively high levels of openness. Openness scores ranged from 2.05 to 7.00, with a mean of 5.06 (median = 5.20). Descriptive statistics, internal consistency estimates and bivariate correlations for all variables are summarized in Table 1.

### Mediation Models

Multiple mediation analyses were conducted using ordinary least squares regression. Indirect effects were tested using nonparametric bootstrapping because indirect effects are often not normally distributed. Bootstrapping was therefore used to obtain more accurate confidence intervals. A total of 5,000 bootstrap resamples were used to estimate bias-corrected 95% confidence intervals for each indirect effect. Indirect effects were considered significant if the confidence interval did not include zero. All variables were standardized prior to analysis so that the resulting coefficients could be interpreted as standardized effects.

| Variable | M | SD | $\alpha$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Wellbeing | 0 | 0.9 | 0.91* | — | | | | | | |
| 2. Conscientiousness | 4.77 | 0.82 | 0.88* | .43** [.40, .47] | — | | | | | |
| 3. Openness | 5.06 | 0.77 | 0.85* | .18** [.13, .22] | .37** [.33, .41] | — | | | | |
| 4. Intrinsic Goal Importance | 5.48 | 1.01 | 0.90* | .24** [.20, .28] | .28** [.24, .32] | .48** [.45, .52] | — | | | |
| 5. Intrinsic Goal Attainment | 4.26 | 1.05 | 0.92* | .57** [.54, .60] | .31** [.27, .35] | .24** [.19, .28] | .46** [.42, .49] | — | | |
| 6. Extrinsic Goal Importance | 3.48 | 1.25 | 0.87* | 0.04 [-.01, .08] | -.06* [-.10, -.01] | -.12** [-.16, -.07] | .23** [.19, .27] | .18** [.14, .22] | — | |
| 7. Extrinsic Goal Attainment | 2.92 | 1.2 | 0.91* | .29** [.25, .33] | -0.01 [-.05, .04] | -.15** [-.20, -.11] | .06** [.02, .11] | .55** [.51, .58] | .66** [.63, .68] | — |

**Table 1.** Summary of Variable Means, Standard Deviations, Internal Consistency, and Bivariate Correlations. M = mean; SD = standard deviation; Cronbach's $\alpha$ = internal consistency of multi-item scales. Values in square brackets represent 95% confidence intervals. The confidence interval indicates a plausible range of population correlations that could have produced the observed sample correlation. *p < .05, **p < .01.

**Figure 1**. Multiple Mediation Model of Conscientiousness on Wellbeing via Intrinsic and Extrinsic Goal Importance. Solid lines indicate significant paths; dotted lines indicate nonsignificant paths. Conscientiousness positively predicted intrinsic and negatively predicted extrinsic goal importance; only intrinsic importance significantly predicted wellbeing, partially mediating the association.

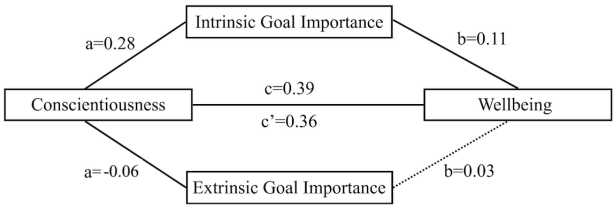
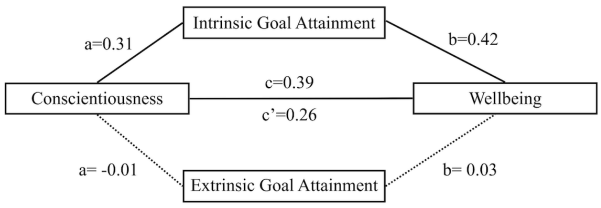



**Figure 2**. Multiple Mediation Model of Conscientiousness on Wellbeing via Intrinsic and Extrinsic Goal Attainment. Solid lines indicate significant paths; dotted lines indicate nonsignificant paths. Conscientiousness positively predicted intrinsic but not extrinsic goal attainment; only intrinsic attainment significantly predicted wellbeing, partially mediating the association.




Four multiple mediation models were conducted in total. For each personality variable (conscientiousness and openness), we ran two models predicting well-being: one testing intrinsic versus extrinsic goal importance as mediators, and the other testing intrinsic versus extrinsic goal achievement as mediators.

A multiple mediation analysis was conducted to examine whether intrinsic goal importance and extrinsic goal importance mediated the relationship between conscientiousness and well-being (Figure 1). All variables were standardized before analysis, so coefficients represent standardized effects. The total effect of conscientiousness on well-being was significant, B=0.39, SE=0.02, t(1890)=20.87, p<.001, 95% CI [0.35,0.42].

When both mediators were entered simultaneously, conscientiousness significantly predicted intrinsic goal importance (B=0.28, SE=0.02, p<.001, 95% CI [0.24,0.32]) and negatively predicted extrinsic goal importance (B=–0.06, SE=0.02, p=.012, 95% CI [–0.11,–0.01]). In turn, intrinsic importance significantly predicted well-being (B=0.11, SE=0.02, p<.001, 95% CI [0.07,0.15]), whereas extrinsic importance was not a significant predictor (B=0.03, SE=0.02, p=.096, 95% CI [–0.01,0.07]).

The indirect effect of conscientiousness on well-being through intrinsic goal importance was significant, B=0.03, 95% bootstrap CI [0.02,0.04], whereas the indirect effect through extrinsic goal importance was not significant, B=0.00, 95% bootstrap CI [–0.01,0.00]. The total indirect effect was B=0.03, 95% bootstrap CI [0.01,0.04]. The direct effect of conscientiousness on well-being remained significant after accounting for both mediators (B=0.36, SE=0.02, p<.001, 95% CI [0.32,0.40]), indicating partial mediation.

We then tested the impact of conscientiousness on well-being via intrinsic versus extrinsic goal attainment using the same type of multiple mediation model (Figure 2). Again, the total effect of conscientiousness on well-being was significant, B=0.39, SE=0.02, t(1890)=20.87, p<.001, 95% CI [0.35,0.43].

When both mediators were entered simultaneously, conscientiousness significantly predicted intrinsic goal attainment (B=0.31, SE=0.02, p<.001, 95% CI [0.27,0.35]) but did not significantly predict extrinsic goal attainment (B=–0.01, SE=0.02, p=.691, 95% CI [–0.05,0.03]). In turn, intrinsic goal attainment significantly predicted well-being (B=0.42, SE=0.02, p<.001, 95% CI [0.38,0.46]), whereas extrinsic goal attainment was not a significant predictor (B=0.03, SE=0.02, p=.096, 95% CI [–0.01,0.07]).

The indirect effect of conscientiousness on well-being through intrinsic goal attainment was significant, B=0.13, 95% bootstrap CI [0.11,0.15], whereas the indirect effect through extrinsic goal attainment was not sig-

nificant, B=0.00, 95% bootstrap CI [0.00,0.00]. The total indirect effect was B=0.13, 95% bootstrap CI [0.10,0.15]. The direct effect of conscientiousness on well-being remained significant after accounting for both mediators (B=0.26, SE=0.02, p<.001, 95% CI [0.22,0.30]), indicating partial mediation.

Using the same type of model, we tested the impact of openness on well-being via intrinsic versus extrinsic goal importance (Figure 3). The total effect of openness on well-being was significant, B=0.16, SE=0.02, t(1890)=7.79, p<.001, 95% CI [0.12,0.20].

When both mediators were entered simultaneously, openness significantly predicted intrinsic goal importance (B=0.48, SE=0.02, p<.001, 95% CI [0.44,0.52]) and negatively predicted extrinsic goal importance (B=–0.12, SE=0.02, p<.001, 95% CI [–0.16,–0.07]). In turn, intrinsic goal importance significantly predicted well-being (B=0.18, SE=0.02, p<.001, 95% CI [0.13,0.23]), whereas extrinsic goal importance was not a significant predictor (B=0.00, SE=0.02, p=.893, 95% CI [–0.04,0.05]).

The indirect effect of openness on well-being through intrinsic goal importance was significant, B=0.09, 95% bootstrap CI [0.06,0.12], whereas the indirect effect through extrinsic goal importance was not significant, B=0.00, 95% bootstrap CI [–0.01,0.01]. The total indirect effect was B=0.09, 95% bootstrap CI [0.06,0.12]. The direct effect of openness on well-being remained significant after accounting for both mediators (B=0.07, SE=0.02, p=.002, 95% CI [0.02,0.12]), indicating partial mediation.

We then tested the impact of openness on well-being via intrinsic versus extrinsic goal attainment using the same type of multiple mediation model (Figure 4). Again, the total effect of openness on well-being was significant, B=0.16, SE=0.02, t(1890)=7.79, p<.001, 95% CI [0.12,0.20].

When both mediators were entered simultaneously, openness significantly predicted intrinsic goal attainment (B=0.24, SE=0.02, p<.001, 95% CI [0.20,0.29]) and extrinsic goal attainment (B=–0.15, SE=0.02, p<.001, 95% CI [–0.19,–0.11]). In turn, intrinsic goal attainment significantly predicted well-being (B=0.52, SE=0.02, p<.001, 95% CI [0.48,0.56]), whereas extrinsic attainment was not a significant predictor (B=–0.02, SE=0.02, p=.336, 95% CI [–0.05,0.02]).

The indirect effect of openness on well-being through intrinsic goal attainment was significant, B=0.12, 95% bootstrap CI [0.10,0.15], whereas the indirect effect through extrinsic goal attainment was not significant, B=0.00, 95% bootstrap CI [0.00,0.01]. The total indirect effect was B=0.13, 95% bootstrap CI [0.10,0.16]. The direct effect of openness on well-being was nonsignificant after accounting for both mediators (B=0.03, SE=0.02, p=.088, 95% CI [–0.00,0.07]), indicating full mediation primarily through intrinsic attainment.

**Figure 3**. Multiple Mediation Model of Openness on Wellbeing via Intrinsic and Extrinsic Goal Importance. Solid lines indicate significant paths; dotted lines indicate nonsignificant paths. Openness positively predicted intrinsic and negatively predicted extrinsic goal importance; only intrinsic importance significantly predicted wellbeing, partially mediating the association.

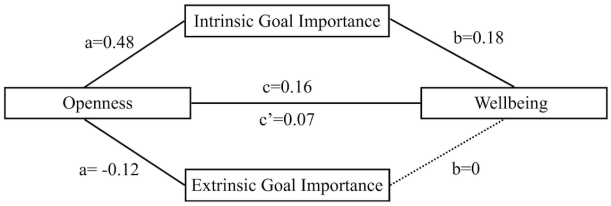
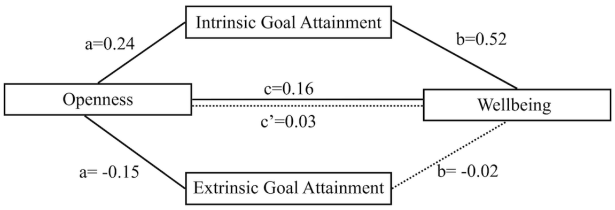



**Figure 4**. Multiple Mediation Model of Openness on Wellbeing via Intrinsic and Extrinsic Goal Attainment. Solid lines indicate significant paths; dotted lines indicate nonsignificant paths. Openness positively predicted intrinsic and negatively predicted extrinsic goal attainment; only intrinsic attainment significantly predicted wellbeing, fully mediating the association.




## Discussion

Both conscientiousness and openness are positively associated with well-being. These personality factors are also positively related to intrinsic goal importance and attainment, which partially mediate their effects on well-being, with the effect of openness on well-being being fully mediated by intrinsic goal attainment.

The direct effect of conscientiousness on well-being is consistent with the literature, which generally finds a positive association. Conscientiousness is linked to greater life satisfaction, more positive affect, and fewer negative emotions. This is largely because conscientious individuals engage in healthier behaviors, experience lower stress, and use more effective coping strategies[24, 25]. They are also more likely to succeed academically, professionally, and financially, which in turn enhances well-being[26, 27]. Longitudinal studies suggest that conscientiousness predicts later well-being, rather than the reverse[25].

Conscientiousness was also a significant predictor of intrinsic goal importance. Conscientious individuals may be more likely to pursue goals for their inherent value because they tend to be organized, disciplined, and motivated by long-term personal growth rather than external rewards. These traits align with valuing intrinsic goals, such as self-development, which provide personal satisfaction. Unexpectedly, conscientiousness also weakly negatively predicted extrinsic goal importance (B=–0.06). Though it was hypothesized conscientiousness would positively predict extrinsic goal aspiration due to increased drive and motivation, this result could potentially be explained because individuals who focus on intrinsic goals dedicate more time and energy to those pursuits, meaning less attention is dedicated to extrinsic goals.

Intrinsic goal aspiration significantly predicted well-being. This aligns with Self-Determination Theory (SDT), which emphasizes that intrinsic aspirations like personal growth, meaningful relationships, and contributing to one's community support the fulfillment of the three basic psychological needs: autonomy, competence, and relatedness[28]. Meeting these needs promotes subjective well-being. Extrinsic goal aspiration did not significantly predict well-being. While extrinsic goals such as financial stability can increase material comfort and indirectly support well-being[29], pursuing them at the expense of intrinsic goals may actually reduce subjective well-being due to unmet psychological needs[28].

The indirect effect of conscientiousness on well-being through intrinsic goal importance was significant, but small (B=0.03), whereas the indirect effect through extrinsic goal importance was not significant. This suggests that intrinsic goal importance partially mediates the relationship between conscientiousness and well-being, though other factors are also involved, including those discussed earlier such as effective coping strategies, health behaviors, and social support.

Conscientiousness significantly predicts intrinsic goal attainment. This is consistent with research showing that people tend to achieve the types of goals to which they aspire[30]. Since conscientious individuals are more likely to aspire to intrinsic goals, they are also more likely to attain them. Additionally, people high in conscientiousness tend to be organized, disciplined, and capable of sticking to plans, which further explains their success in achieving these goals. In contrast, conscientiousness did not significantly predict extrinsic goal attainment, which was unexpected. A possible explanation for this finding is that conscientious individuals tend to prioritize intrinsic over extrinsic goals, making them less likely to pursue or attain extrinsic goals.

Intrinsic goal attainment significantly predicted well-being, even more strongly than intrinsic aspiration. This makes sense because achieving intrinsic goals directly satisfies basic psychological needs and provides tangible evidence of competence, autonomy, and relatedness, resulting in a stronger boost to well-being[28, 30]. Extrinsic goal attainment, however, was not associated with well-being, supporting the idea that while certain aspects of extrinsic goal attainment may contribute to well-being, achieving extrinsic goals at the expense of intrinsic ones may undermine these benefits.

The indirect effect of conscientiousness on well-being through intrinsic attainment was significant, whereas the indirect effect through extrinsic attainment was not. This indirect effect was also larger than the effect of conscientiousness on well-being through intrinsic goal aspiration, indicating that intrinsic goal attainment partially explains the relationship between conscientiousness and well-being. Because conscientious individuals are more likely to attain intrinsic goals, they are more likely to experience greater well-being. However, this mediation is only partial, suggesting that other factors also contribute to the relationship between conscientiousness and well-being.

Openness significantly predicts well-being, though to a lesser extent than conscientiousness. Openness also significantly positively predicts intrinsic goal aspiration. Individuals high in openness are more likely to value personal growth, self-development, and meaningful experiences, which are all core aspects of intrinsic goals. Their curiosity and interest in new experiences make them more inclined to set goals that provide internal satisfaction rather than external rewards[2, 9]. Openness also negatively predicted extrinsic goal aspiration. Because individuals high in openness tend to prioritize intrinsic goals that reflect their values, they may intentionally deemphasize extrinsic goals in favor of pursuits that satisfy personal growth.

The indirect effect of openness on well-being through intrinsic goal importance was significant, whereas the indirect effect through extrinsic goal importance was not. One possible explanation for this is that people high in openness tend to experience greater well-being partly because they value intrinsic goals. This once again aligns with Self-Determination Theory, which emphasizes that pursuing intrinsic aspirations supports the fulfillment of three basic psychological needs[28]. However, the relationship between openness and well-being remains significant even when accounting for intrinsic and extrinsic goal aspirations, indicating that other factors likely play a role. For example, openness may also promote flexibility in thinking and problem-solving, which helps individuals cope with challenges and reduce stress[31].

Openness significantly predicts intrinsic goal attainment, though less strongly than conscientiousness. This could be because, while people high in openness often strive for intrinsic goals, conscientious individuals are more likely to achieve their goals due to stronger follow-through and self-discipline. Although it was not initially expected that people higher in openness would be more likely to attain their goals, this makes sense when considering that they may pursue intrinsic goals more frequently, which would naturally increase their likelihood of success. Openness also negatively predicts extrinsic goal attainment, which could be because individuals high in openness are less motivated by extrinsic goals.

As discussed earlier, intrinsic goal attainment significantly predicts well-being, and the indirect effect of openness on well-being through intrinsic attainment was significant (B=0.12), whereas the indirect effect through extrinsic attainment was not. In fact, the relationship between openness and well-being was fully mediated by intrinsic goal attainment. This would support the notion that the main reason people high in openness experience greater well-being is because they are achieving intrinsic goals.

These findings have important implications for understanding how personality relates to well-being. They highlight one mechanism through which traits like conscientiousness and openness can promote well-being: conscientiousness fosters intrinsic goal aspirations and the discipline to achieve them, while openness encourages the pursuit of intrinsic goals that align with personal values.

A limitation of this study is the way well-being was operationalized. Here, well-being was measured using positive affect, negative affect, and subjective life satisfaction. Although this is well-validated, it may not capture other dimensions, such as physical health or broader functioning. Moreover, distinctions between hedonic and eudaemonic well-being were not assessed, which might have yielded different results. Depending on the definition of well-being, goal aspirations and attainment could relate differently to personality traits. The sample was mostly from the United States (83%), with the remainder from 50 other countries; due to cultural, economic, and personality differences, the findings may generalize best to U.S.-based online populations with higher openness and conscientiousness. It is important to note that causal claims cannot be drawn from these correlational data. Other variables, such as socioeconomic status, life stress and social support, may contribute to the observed relationships, highlighting the need for caution in interpreting these effects as causal. Future research could incorporate additional measures of well-being, including health outcomes, and examine how other personality traits, such as neuroticism, relate both to the goals people pursue and to their overall well-being. This would provide a more complete understanding of how personality influences well-being.

## Conclusion

This study demonstrates that personality traits, specifically conscientiousness and openness, contribute to well-being in part through goal setting and attainment. Conscientious individuals are more likely to pursue and achieve intrinsic goals, which in turn enhances feelings of autonomy, competence, and relatedness, supporting well-being. Openness similarly promotes intrinsic goal aspirations and attainment, although to a lesser extent, and de-emphasizes extrinsic goals. Intrinsic goals consistently emerged as a stronger predictor of well-being than extrinsic goals, highlighting the importance of pursuing personally meaningful objectives. Overall, these findings highlight the role of intrinsic goal pursuit as a key mechanism linking personality to well-being, in line with the predictions of Self-Determination Theory. While conscientiousness and openness contribute to well-being partly through goal-related pathways, it is evident other mechanisms also contribute to these relationships. Future research should examine additional personality traits, incorporate broader measures of well-being, and further explore how different types of goals mediate the relationship between personality and well-being.

Acknowledgements

I would like to thank my supervisor, Dr. William Ryan, for his guidance and support throughout this project. I am also grateful to the WISH (Well-being, Identity, Stigma, and Health) Lab at the University of Toronto for their ongoing support and constructive feedback throughout the research process.

Data Availability Statement

The data that support the findings of this study are not publicly available due to participant privacy.

## References

1. Baumert, A. Integrating personality structure, personality process, and personality development. *Eur. J. Pers* **31,** 503–528 (2017). 10.1002/per.2115
2. McCrae, R. & Costa Jr, P. Personality trait structure as a human universal. *Am. Psychol* **52,** 509–516 (1997). 10.1037/0003-066X.52.5.509
3. Roberts, B. *Conscientiousness. in Handbook of individual differences in social behavior* (eds Leary, M. & Hoyle, R.) Guilford Press, 2009).
4. Anglim, J. & Grant, S. Predicting psychological and subjective well-being from personality: Incremental prediction from 30 facets over the Big 5. *J. Happiness Stud* **17,** 59–80 (2016). 10.1007/s10902-014-9583-7
5. Fuente, J. et al. The Big Five factors as differential predictors of self-regulation, achievement emotions, coping and health behavior in undergraduate students. *BMC Psychol* **12, Article 267** (2024). 10.1186/s40359-024-01768-9
6. Atherton, O., Robins, R., Rentfrow, P. & Lamb, M. Personality correlates of risky health outcomes: Findings from a large Internet study. *J. Res. Pers* **50,** 56–60 (2014). 10.1016/j.jrp.2014.03.002
7. Gartland, N., O'Connor, D., Lawton, R. & Ferguson, E. Investigating the effects of conscientiousness on daily stress, affect and physical symptom processes: A daily diary study. *Br. J. Health Psychol* **19,** 311–328 (2014). 10.1111/bjhp.12077
8. Javaras, K. et al. Conscientiousness predicts greater recovery from negative emotion. *Emotion* **12,** 875–881 (2012). 10.1037/a0028105
9. Judge, T. & Ilies, R. Relationship of personality to performance motivation: A meta-analytic review. *J. Appl. Psychol* **87,** 797–807 (2002). 10.1037/0021-9010.87.4.797
10. Roberts, B., Lejuez, C., Krueger, R., Richards, J. & Hill, P. What is conscientiousness and how can it be assessed? Dev. *Psychol* **50,** 1315–1330 (2014). 10.1037/a0031109
11. Bradshaw, E., Ryan, R. & Deci, E. A meta-analysis of the dark side of the American dream: Evidence for the universal wellness costs of prioritizing extrinsic over intrinsic goals. *J. Pers. Soc. Psychol* **124,** 873–899 (2023). 10.1037/pspp0000431
12. Schmuck, P., Kasser, T. & Ryan, R. Intrinsic and extrinsic goals: Their structure and relationship to well-being in German and U.S. college students. *Soc. Indic. Res* **50,** 225–241 (2000). 10.1023/A:1007084005278
13. Sheldon, K. & Elliot, A. Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *J. Pers. Soc. Psychol* **76,** 482–497 (1999). 10.1037/0022-3514.76.3.482
14. Lodi-Smith, J. & Roberts, B. Social investment and personality: A meta-analysis of the relationship of personality traits to investment in work, family, religion, and volunteerism. *Pers. Soc. Psychol. Rev* **11,** 68–86 (2007). 10.1177/1088868306294590
15. DeYoung, C., Quilty, L. & Peterson, J. Between facets and domains: 10 aspects of the Big Five. *J. Pers. Soc. Psychol* **93,** 880–896 (2007). 10.1037/0022-3514.93.5.880
16. Kasser, T. & Ryan, R. A dark side of the American Dream: Correlates of financial success as a central life aspiration. *Pers. Soc. Psychol* **65,** 410–422 (1993). 10.1037/0022-3514.65.2.410
17. Diener, E. et al. New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Soc. Indic. Res* **97,** 143–156 (2010). 10.1007/s11205-009-9493-y
18. Diener, E., Emmons, R., Larsen, R. & Griffin, S. The Satisfaction With Life Scale. *J. Pers. Assess* **49,** 71–75 (1985). 10.1207/s15327752jpa4901_13
19. Hallford, D. & Mellor, D. Development and validation of the Awareness of Narrative Identity Questionnaire (ANIQ. *Assessment* **24,** 399–413 (2017). 10.1177/1073191115607046
20. Chen, B. et al. Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motiv. Emot* **39,** 216–236 (2015). 10.1007/s11031-014-9450-1
21. Ware, J. & Sherbourne, C. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Med. Care* **30,** 473–483 (1992). 10.1097/00005650-199206000-00002
22. Pennebaker, J. *The Pennebaker Inventory of Limbic Languidness (PILL* Unpublished inventory. 1982.
23. Condon, D. & Revelle, W. The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence* **43,** 52–64 (2014). 10.1016/j.intell.2014.01.004
24. Bogg, T. & Roberts, B. The case for conscientiousness: Evidence and implications for a personality trait marker of health and longevity. *Ann. Behav. Med* **45,** 278–288 (2013). 10.1007/s12160-012-9454-6
25. Hu, Y., Wang, Z. & Fan, Q. The relationship between conscientiousness and well-being among Chinese undergraduate students: A cross-lagged study. *Int. J. Environ. Res. Public Health* **19,** 13565 (2022). 10.3390/ijerph192013565
26. Duckworth, A., Weir, D., Tsukayama, E. & Kwok, D. Who does well in life? Conscientious adults excel in both objective and subjective success. *Front. Psychol* **3,** 356 (2012). 10.3389/fpsyg.2012.00356
27. Chen, S., Cheung, A. & Zeng, Z. Big Five personality traits and university students academic performance: A meta-analysis. *Personality Individ. Differ* **240,** 113163 (2025). 10.1016/j.paid.2025.113163
28. Deci, E. & Ryan, R. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol* **55,** 68–78 (2000). 10.1037/0003-066X.55.1.68
29. Thomson, R. et al. How do income changes impact on mental health and wellbeing for working-age adults? A systematic review and meta-analysis. *Lancet Public Health* **7,** 515–528 (2022). 10.1016/S2468-2667(22)00058-5
30. Sheldon, K. & Elliot, A. Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *J. Pers. Soc. Psychol* **76,** 482–497 (1999). 10.1037/0022-3514.76.3.482
31. Chen, L., Qu, L. & Hong, R. Pathways linking the Big Five to psychological distress: Exploring the mediating roles of stress mindset and coping flexibility. *J. Clin. Med* **11,** 2272 (2022). 10.3390/jcm11092272

Submitted: 01/09/2026
Accepted: 02/20/2026
Published: 04/10/2026

Review Article

[1]Department of Physiology, McGill University, Montreal, QC, Canada

Keywords

cerebrospinal fluid, acid–base, Stewart approach, strong ion difference, choroid plexus, blood–CSF barrier, neurocritical care, ventilation

Email Correspondence

shenglin.qu@mail.mcgill.ca

https://doi.org/10.26443/msurj.v21i1.400



Qu, Samuel[1]

# Cerebrospinal Fluid Acid-Base Physiology in Critical Illness: A Stewart Physicochemical Interpretation

## Abstract

Cerebrospinal fluid (CSF) constitutes the physicochemical milieu bathing the brain and brainstem and can diverge from arterial blood in both magnitude and time course during critical illness. This review synthesizes human and experimental evidence relevant to a Stewart physicochemical interpretation of CSF acid-base physiology, emphasizing compartmental kinetics, blood–CSF barrier and choroid plexus transport, and methodological constraints that complicate measurement. Because CSF contains minimal weak-acid buffering under physiological conditions, CSF pH is often dominated by $PCO_2$ and strong ion difference (SID), making CSF a tractable compartment for compartment-aware physicochemical reasoning. Classical clinical observations and modern simultaneous CSF–arterial datasets in pregnancy and aneurysmal subarachnoid hemorrhage illustrate paired shifts in SID and $PCO_2$ that are not inferable from arterial blood gases alone. Transporter-dependent recovery from hypercapnia, pH-sensitive ion fluxes, and chloride-linked transport perturbation studies support that CSF acid-base homeostasis is actively regulated and plausibly vulnerable in neuroinflammation and brain injury. We outline ICU implications, focusing on why arterial normalization may not ensure central normalization and why time-resolved ventricular datasets are the key prerequisite for bedside translation.

## Introduction

CSF is not merely a passive ultrafiltrate of plasma. It is the physicochemical environment bathing the brain and brainstem and therefore sits at the interface of two ICU priorities: ventilatory management and neurological physiology. In critically ill patients, particularly those with acute brain injury, clinicians routinely manipulate ventilation to control arterial pH and $PaCO_2$. Yet the therapeutic targets that matter for central chemosensitivity and cerebrovascular tone are more plausibly CSF and brain extracellular chemistry, which can diverge from arterial blood in both magnitude and time course[1].

The mechanistic basis for this dissociation is kinetic and compartmental. $CO_2$ crosses the blood–brain and blood–CSF interfaces rapidly, whereas ionic and other non-volatile components (for example chloride, lactate, and bicarbonate-equivalent handling) are constrained by barrier transport and therefore equilibrate more slowly and often incompletely. Clinically important mismatch states can follow: arterial blood gases appear corrected while CSF chemistry remains displaced (or vice versa), with potential consequences for respiratory drive and neurovascular regulation[1]. Classic clinical observations in systemic metabolic derangements similarly show that CSF acid-base relationships are not reliably inferable from plasma alone[2].

Pre-Stewart clinical work already anticipated the clinical relevance of the central compartment. Serial paired blood–CSF observations in severe systemic acid-base disorders showed that neurological status correlated more closely with CSF pH than with arterial acidemia alone, and that blood and CSF can evolve on different trajectories during illness and treatment[3]. These observations align with the ICU reality that arterial variables are readily manipulated, while central variables are coupled through barriers with distinct kinetics.

Stewart's physicochemical framework offers a coherent language for this compartmental problem because CSF composition differs from plasma. Under physiological conditions CSF has very low protein and weak-acid buffering, so CSF pH is often dominated by $PCO_2$ and strong ion difference (SID), with total weak non-volatile acids (ATOT) playing a smaller role than in blood[1]. This simplification is meaningful in contrast to plasma, where weak-acid (protein) charge chemistry complicates modeling and interpretation[4]. Within Stewart's formulation:

$$pH = f(\mathrm{PCO_2},\ \mathrm{SID},\ \mathrm{ATOT}), \tag{1}$$

and a pragmatic operational definition of CSF SID is:

$$\mathrm{SID_{CSF}} \approx \left([\mathrm{Na^+}] + [\mathrm{K^+}] + [\mathrm{Ca^{2+}}] + [\mathrm{Mg^{2+}}]\right) - \left([\mathrm{Cl^-}] + [\mathrm{Lactate^-}] + \ldots\right) \tag{2}$$

Because ATOT is typically small in normal CSF, bicarbonate is best treated as a dependent readout of electroneutrality and $CO_2$ chemistry rather than a controlling variable. This perspective is consistent with Stewart-quantitative development and clinical translation in critical care[5,6,7].

Recent simultaneous CSF–arterial studies provide clinically relevant constraints. In late pregnancy, sustained hypocapnia is accompanied by coordinated changes in CSF and plasma physicochemical variables and altered CSF–plasma gradients, consistent with compartment-specific setpoints and adaptation rather than simple arterial mirroring[8]. In neurocritical care, spontaneously breathing patients with aneurysmal subarachnoid hemorrhage (SAH) show CSF profiles interpretable as an interaction between lactate-associated SID shifts and CSF $CO_2$ dynamics, including a SID-lowering burden that can coexist with low CSF $PCO_2$ and near-normal CSF pH[9].

A Stewart interpretation must also be grounded in barrier biology. The choroid plexus and blood–CSF barrier express transport systems capable of shaping CSF composition and recovery kinetics. Experimental studies identify transporter-dependent contributions ranging from pH-sensitive $Na^+/H^+$ exchange-related fluxes during metabolic disturbances to bicarbonate-related transport required for recovery from

hypercapnia[10,11]. This review therefore aims to synthesize human and experimental evidence relevant to a Stewart interpretation of CSF acid-base physiology and to outline translational implications for ICU practice, including why ABG normalization may not guarantee central normalization and why higher-temporal-resolution CSF datasets are needed for bedside translation[1,12].

## Methods

This is a focused narrative review integrating (i) human studies reporting CSF acid-base variables with sufficient data to interpret physicochemical determinants (including simultaneous CSF–arterial sampling in pregnancy and neurocritical illness)[8,9], (ii) classic clinical physiology in systemic metabolic derangements that illustrates blood–CSF dissociation[2,3,13], and (iii) mechanistic studies of choroid plexus and blood–CSF barrier transport relevant to CSF pH regulation and strong-ion composition[10,11,14,15,16]. We also cite integrative sources on cerebrovascular $CO_2$ responsiveness, chemoreflex physiology, and weak-acid modeling where they sharpen compartmental interpretation[4,17,18,19]. From each eligible study we extracted clinical/model context, sampling site (lumbar vs cisternal vs ventricular), measured variables (pH, $PCO_2$, electrolytes and lactate when available), and sampling/measurement methods, with particular attention to pre-analytical vulnerabilities of CSF pH/$PCO_2$[12]. Where feasible we interpreted coupling using gradients:

$$\Delta PCO_2 = PCO_2(\text{CSF}) - PCO_2(\text{arterial}) \tag{3}$$

and

$$\Delta SID = SID(\text{CSF}) - SID(\text{plasma}). \tag{4}$$

Because the goal is mechanistic integration and ICU framing rather than pooled effect estimates, no meta-analysis was attempted.

## Physicochemical framework for CSF Acid-Base Interpretation

In ICU practice, bicarbonate-centered descriptors (bicarbonate, base excess, $PaCO_2$) remain clinically useful, but compartmental problems are often easier to reason about in terms of independent determinants and coupling time scales. Stewart's framework treats pH as determined by $PCO_2$, SID, and ATOT. In normal CSF, ATOT is typically small because protein buffering is minimal, which makes CSF a comparatively transparent compartment for physicochemical reasoning[1]. This is best understood relative to plasma, where weak-acid chemistry is more complex and protein charge modeling can materially affect pH interpretation[4].

Operationally, SID is an accounting of strong cations minus strong anions. In CSF, lactate and chloride are often the most clinically relevant strong anions, and their inclusion is necessary for interpretable SID calculations in pathology[9]. A practical advantage of this framework is that it makes gradients interpretable: $\Delta PCO_2$ and $\Delta SID$ can be viewed as state variables that reflect blood–CSF coupling and transport-limited adaptation rather than as noise.

Because much transporter physiology is reported in bicarbonate/$H^+$-transport language, a consistent mapping is needed. In this review, we describe transporter findings in their native terms but interpret their acid-base consequences through changes in the strong-ion environment and electroneutrality, with bicarbonate treated as a dependent outcome. This avoids mixing causal and dependent variables and aligns with Stewart-quantitative and bedside physicochemical traditions[5,6,7].

## Measurement Validity and Compartmental Constraints

CSF pH and $PCO_2$ are unusually sensitive to sampling artefacts because CSF has minimal non-volatile buffering. The dominant confounder is $CO_2$ loss during sampling and transfer, which increases measured pH and decreases measured $PCO_2$. Davies' *in vivo*/direct comparison demonstrated systematic, non-trivial error introduced by syringe technique, with $PCO_2$ underestimation increasing as true $PCO_2$ rises[12]. Such errors can materially distort small blood–CSF gradients and complicate inference about coupling.

Methodological rigor therefore becomes part of interpretation. Some classic studies provide relatively explicit sampling details; for example, Ohman *et al* reported anaerobic collection, cold handling, and rapid analysis in diabetic ketoacidosis, improving confidence that discordant blood–CSF trajectories reflect physiology rather than dominant artefact[13]. Contemporary studies should report anaerobic handling, time-to-analysis, temperature correction, and timing relative to changes in ventilation or drainage.

Sampling site is also a biological variable. Lumbar, cisternal, and ventricular CSF reflect different flow and exchange contexts; neurocritical studies frequently use ventricular access via external ventricular drains, whereas much older physiology is lumbar-based. Without explicit stratification by site and drainage conditions, apparent disease effects may partly reflect compartmental heterogeneity[9]. Finally, analytic completeness matters: lactate omission biases SID upward, and inconsistent inclusion of $Ca^{2+}/Mg^{2+}$ or total vs ionized values limits cross-study comparability[9].

The remaining major gap is temporal resolution. Much human CSF literature is cross-sectional, yet the physiology is dynamic: $CO_2$ equilibrates rapidly, while strong-ion/strong-anion adjustment is transport-limited and slower[1]. Continuous *in vivo* CSF pH recording has been demonstrated experimentally and resolves recovery kinetics during hypercapnia and transporter perturbation[11], but analogous high-temporal-resolution human ICU datasets remain scarce.

## Human and Disease Evidence: Coupling Patterns that Matter Clinically

A Stewart-oriented reading is most useful when it asks where the perturbation originates and how coupling unfolds over time. Classic clinical studies already show that blood and CSF can behave differently in severe systemic disorders. Posner and Plum demonstrated that neurological status correlated more closely with CSF pH than with arterial acidemia alone, supporting a central-compartment focus in interpreting neurologic consequences of systemic acid-base disturbances[3]. Ohman *et al* further showed discordant blood–CSF trajectories during treatment of diabetic ketoacidosis, including blood pH improvement while CSF pH fell in early treatment in several cases, alongside persistent differences in $PCO_2$ and bicarbonate and slower equilibration of glucose and ketoacids[13]. These pre-Stewart observations are consistent with barrier-limited coupling and non-identical time constants.

Pregnancy provides a relatively clean human model of sustained hypocapnia. Simultaneous sampling in late pregnancy shows coordinated adaptation across compartments with altered CSF–plasma gradients rather than simple arterial mirroring[8]. From an ICU perspective, the transferable principle is that sustained ventilatory shifts can drive compartment-specific setpoints and transport-limited adaptation, and that gradients may be informative descriptors of coupling (Figure 1; Figure 2).

In neurocritical illness, SAH illustrates a clinically instructive trade-off pat-

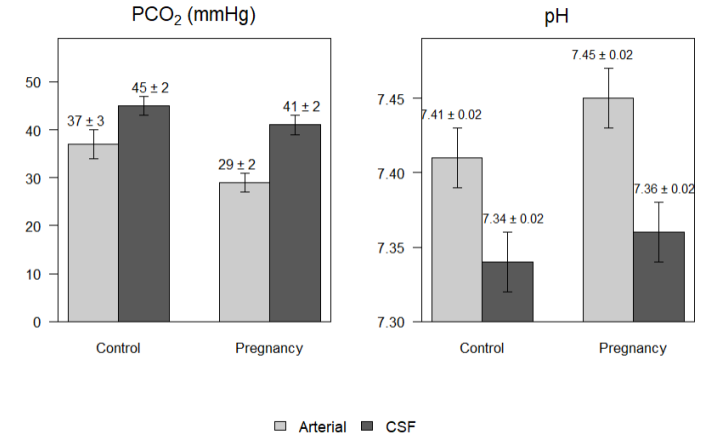



**Figure 1.** Arterial and cerebrospinal fluid acid–base variables in third-trimester pregnancy and non-pregnant controls. Bars show mean (SD) of simultaneously collected arterial blood and CSF samples obtained at the time of spinal anesthesia (controls, $n = 13$; pregnancy, $n = 20$). Variables shown are arterial and CSF pH (dimensionless) and $PCO_2$ (mmHg). Values are extracted from Zadek *et al.* (Br J Anaesth, 2022), Tables 1–2.

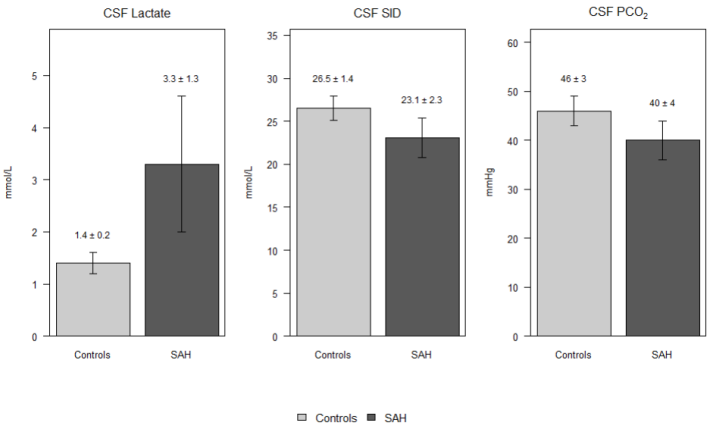



**Figure 2.** Blood–CSF coupling gradients in third-trimester pregnancy and non-pregnant controls. Bars show mean (SD) of $\Delta PCO_2$ and $\Delta SID$ computed as CSF minus plasma values from simultaneously collected samples obtained at the time of spinal anesthesia (controls, $n = 13$; pregnancy, $n = 20$). $\Delta PCO_2$ is expressed in mmHg. $\Delta SID$ is expressed in mM and is negative when plasma SID exceeds CSF SID. Values are extracted from Zadek *et al.* (Br J Anaesth, 2022), Table 2.

tern. In spontaneously breathing SAH patients, CSF SID is reduced largely due to higher CSF lactate, while CSF $PCO_2$ is also reduced; CSF pH may remain near normal and ATOT does not differ materially from controls[9]. Interpreted physicochemically, a SID-lowering strong-anion burden coexists with hypocapnia that counterbalances acidification (Figure 3). This cautions against treating normal CSF pH as evidence of metabolic neutrality and highlights why ABG-only reasoning can miss clinically relevant central states[1,9].

Classic systemic metabolic acidoses further support a strong-anion burden framing. Marks *et al* reported a strong inverse relationship between CSF bicarbonate and total CSF organic acids[2]:

$$\left[HCO_3^-\right]_{CSF} = 13.0 - 0.82\,(\text{total organic acids}),$$
$$r = -0.93, \quad p < 0.001. \tag{5}$$

This relationship supports interpretation of CSF bicarbonate as a dependent consequence of organic acid burden rather than a primary controller. Together, the DKA datasets[13] and organic-acid relationship[2] provide a consistent narrative: coupling is constrained by barriers, and ionic/strong-anion components may persist after arterial correction.

Infectious CNS disease remains relevant as a high-metabolic-load translational scenario. Experimental meningitis produces major derangements in respiration and circulation consistent with a central chemical environment capable of influencing ventilatory output[20]. Although older datasets often lack complete ion panels required for full Stewart decomposition, they emphasize that CNS pathology can alter CSF chemistry, barrier properties, and coupling in clinically meaningful ways.

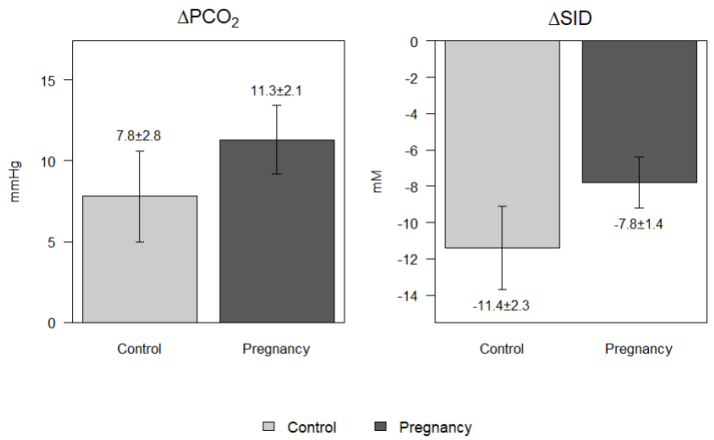



**Figure 3.** Cerebrospinal fluid strong-anion/strong-ion and $CO_2$-related variables in aneurysmal subarachnoid hemorrhage and controls. Bars show mean (SD) of ventricular CSF samples collected simultaneously with arterial blood in spontaneously breathing SAH patients ($n = 20$) and control subjects undergoing elective surgery under spinal anesthesia ($n = 25$). Variables shown are CSF lactate (mmol/L), CSF strong ion difference (SID, mmol/L), and CSF $PCO_2$ (mmHg). Values are extracted from Langer *et al.* (Neurocrit Care, 2022), Table 1.

Across these settings, a useful synthesis is to classify disturbances along two axes: perturbation origin (systemic/peripheral vs central/local) and propagation mechanism ($CO_2$-dominant rapid diffusion vs transport-limited ionic/strong-anion adjustment). This avoids over-rigid phenotype labels while retaining the clinically relevant kinetic asymmetry.

## Mechanistic Substrate: Transport, SID Shaping, and Recovery Kinetics

A Stewart interpretation gains translational traction only if independent determinants map to mechanisms. In CSF, two levers dominate: fast $CO_2$ coupling and transport-limited shaping of the strong-ion environment. The choroid plexus and blood–CSF barrier regulate CSF composition through coordinated transport of bicarbonate equivalents and strong ions, particularly $Na^+$ and $Cl^-$[1,14].

Evidence for pH-sensitive $Na^+$ flux pathways consistent with $Na^+/H^+$ exchange comes from Murphy and Johanson, who demonstrated pH-dependent $Na^+$ flux with an amiloride-inhibitable component during acute metabolic disturbances[10]. In Stewart terms, such fluxes contribute to transport-limited reorganization of the strong-ion environment rather than passive buffering alone.

Ion-transport perturbation studies provide additional mechanistic anchors that are directly interpretable in strong-ion terms. Johnson, Frankel, and Kazemi showed that during hypercapnia, cerebroventricular (but not intravenous) furosemide altered the expected CSF chloride–bicarbonate pattern, producing a greater fall in CSF chloride without the usual rise in CSF bicarbonate during $CO_2$ exposure[16]. Although reported in bicarbonate/chloride language, the key implication is that furosemide-sensitive transport processes can reshape the strong-ion environment during $CO_2$ loading and thereby shape the dependent bicarbonate response.

Recovery kinetics during hypercapnia provide a causal link from trans-

porter capacity to CSF pH resilience. Christensen *et al* used continuous *in vivo* CSF pH recording and demonstrated that disruption of the choroid plexus sodium–bicarbonate cotransporter NBCe2 markedly attenuates CSF pH recovery during sustained hypercapnia[11]. This finding is particularly important for ICU translation because it implies that the same ventilatory perturbation could yield different central trajectories depending on barrier transport capacity.

Broader epithelial machinery plausibly contributes to disease-dependent kinetics. The expression of luminal membrane acid-extrusion related components such as ClC-7 and NHE6 in murine choroid plexus strengthens biological plausibility for an active epithelial toolkit that could be modulated by inflammation or injury[15]. Direct quantitative mapping from specific components to bedside SID trajectories remains limited, but the existence of this machinery supports the view that CSF acid-base homeostasis is actively regulated and therefore vulnerable under pathology.

## Ventilatory Control and Chemosensitivity: Supported Inferences and Boundaries

The translational promise of applying Stewart to CSF is a clearer bridge from chemistry to ventilatory control. Central chemoreception responds to the chemical environment of the CNS/CSF rather than arterial bicarbonate per se[1,5,6]. Within a Stewart framework, strong ions influence ventilation indirectly by altering $[H^+]$ via SID at a given $PCO_2$, with chemosensitive networks transducing the resulting chemical signal.

A physiologically realistic interpretation must also acknowledge integration of central and peripheral chemoreflex mechanisms. The carotid body provides major peripheral input via multimodal sensing of blood $O_2$, $CO_2$, pH, and metabolites including lactate[18]. This boundary condition matters clinically because ventilatory output reflects integrated drives, and central compartment chemistry is a key component but not necessarily the sole determinant in all contexts.

Within this integrated view, SAH provides a compelling example of why pH alone can mislead. A lactate-associated SID-lowering burden can coexist with hypocapnia and near-normal CSF pH[9]. Clinically, ABG-only reasoning can miss this paired central state, and $PaCO_2$ manipulation could shift central pH rapidly before transport-limited SID adaptation occurs[1,9].

Experimental physiology supports plausibility that ionic state can alter ventilatory control. In potassium depletion, Nattie and Tenney demonstrated altered control of breathing in awake rats[21]. Although not framed in Stewart terms and not providing full CSF strong-ion panels, this work supports the general principle that electrolyte states can modulate ventilatory regulation, motivating more direct tests that isolate SID changes at controlled $CO_2$.

Modern circuit-level work, particularly on the retrotrapezoid nucleus (RTN), has advanced understanding of central $CO_2/H^+$ chemoreception and emphasizes multicellular contributions to chemosensory function[19]. However, current circuit-level evidence remains primarily $CO_2/H^+$-centric and does not yet provide direct tests that isolated CSF SID manipulation at constant $CO_2$ alters defined circuit output *in vivo*, particularly in humans. Circuit-specific claims about SID modulation should therefore remain hypothesis-level.

## ICU Implications: Toward Compartment-Aware Interpretation

The ICU relevance of a CSF Stewart framework is not to replace ABG analysis, but to clarify when ABG normalization is an incomplete proxy for the central physicochemical environment governing respiratory drive and potentially cerebrovascular physiology[19]. Because CSF has minimal weak-acid buffering and is shaped by barrier transport with distinct time constants, CSF pH can be maintained near normal by hypocapnia despite a substantial SID-lowering burden, and conversely CSF may remain displaced despite apparent arterial correction.

A practical prediction is kinetic: $PaCO_2$ changes transmit quickly to CSF, while transport-limited ionic adaptation is slower[1]. In patients who already carry a SID-lowering burden, such as lactate-associated changes in neurocritical illness[9], $PaCO_2$ adjustment may shift central pH rapidly before meaningful strong-ion compensation develops. This reframes the clinical question from "is $PaCO_2$ corrected" to "what is the paired central state ($PCO_2$, SID), and how will an intervention move it over time."

$PaCO_2$ manipulation also has parallel cerebrovascular consequences. Human data show nonlinear cerebrovascular responses to $CO_2$ and pressure-coupled effects at higher $CO_2$ ranges, reinforcing that ventilatory interventions can simultaneously alter central chemistry and cerebral hemodynamics[17]. This strengthens the case for cautious, compartment-aware interpretation rather than ABG-only reasoning.

Mechanistic studies further imply potential heterogeneity in central recovery capacity. $NBCe_2$-dependent recovery during hypercapnia[11] and pH-sensitive $Na^+$ flux pathways consistent with exchanger activity[10] support the idea that barrier transport capacity can determine the time constant and completeness of central compensation. In neuroinflammation, hemorrhage, or injury, transporter function and permeability may change, plausibly contributing to variable CSF responses to similar arterial interventions.

A clinically usable CSF Stewart bedside framework would minimally require standardized sampling to minimize $CO_2$ loss and enable reliable gradients[12], a core chemistry panel enabling SID computation ($Na^+$, $K^+$, $Cl^-$, lactate $\pm$ $Ca^{2+}/Mg^{2+}$), ideally paired with simultaneous arterial sampling[8,9], and dynamic data rather than snapshots. Experimental proof of principle exists for continuous CSF pH recording[11]. Consistent with ICU bedside physicochemical pedagogy[22], clinically usable outputs would likely be simple descriptors such as $\Delta PCO_2(t)$, $\Delta SID(t)$, and their relation to CSF pH trajectories, rather than opaque multi-equation summaries.

## Limitations and Gaps

Despite conceptual clarity, the literature remains constrained by methodological heterogeneity, limited temporal resolution, and incomplete mechanistic integration.

First, pre-analytical and analytical heterogeneity limits cross-study comparability. $CO_2$ loss during sampling can materially bias CSF pH and $PCO_2$ and thereby distort gradients and coupling inference[12]. Variation in measured ion panels and SID definitions further limits comparability, especially if lactate is omitted[9].

Second, compartmental heterogeneity matters. Lumbar, cisternal, and ventricular compartments are not interchangeable; drainage conditions and local metabolism can shift composition. ICU translation often relies on ventricular access, whereas many historical datasets are lumbar-based[9].

Third, most human evidence remains cross-sectional. Yet time con-

stants are central to CSF Stewart interpretation: $CO_2$ equilibrates rapidly, while strong-ion/strong-anion adjustment is transport-limited and slower[1]. High-temporal-resolution human datasets that couple ventilator changes to CSF pH/$PCO_2$ and electrolytes remain rare, and this is a major barrier to bedside translation.

Fourth, mechanistic integration under ICU-relevant pathologies is incomplete. Transporter dependence is established for recovery from hypercapnia[11], exchanger-like fluxes are demonstrable in metabolic disturbances[10], and chloride-linked transport perturbation studies support transport-regulated ionic rearrangement during $CO_2$ loading[16]. However, direct linking of specific critical illness pathologies to quantitative changes in transporter function and CSF SID trajectories remains limited[15].

Fifth, ventilatory-control circuitry integration in Stewart terms is undertested. Modern RTN-centered work advances $CO_2/H^+$ circuit models[19], but direct tests isolating SID changes at controlled $CO_2$ with parallel circuit readouts remain scarce. Thus, SID-specific circuit claims should remain hypothesis-level.

Finally, language accessibility and selection bias remain practical limitations. Relevant historical work exists in non-English sources (including German-language physiology)[23]. English-only synthesis may therefore underrepresent certain early observations. In addition, although CSF is often low-ATOT under physiological conditions, this simplification may weaken when CSF protein increases, and plasma-side weak-acid complexity provides a cautionary parallel[4].

## Future Directions

Four priorities follow directly from the current evidence base. First, high-temporal-resolution human ventriculostomy datasets are needed, coupling ventilator changes to CSF pH, $PCO_2$, electrolytes, lactate, and paired arterial sampling. Second, studies should connect ICU-relevant pathologies (SAH, TBI, neuroinflammation) to barrier transport capacity and CSF physicochemical trajectories using chemistry plus transporter biomarkers or physiologic surrogates. Third, integration with chemosensory circuitry should test whether Stewart-defined ionic state variables alter ventilatory response curves or circuit output at controlled $CO_2$, leveraging modern RTN/network paradigms[19]. Fourth, broader inclusion of non-English historical literature may improve continuity of mechanistic synthesis[23].

## Conclusion

CSF acid-base physiology in critical illness is best understood as a compartmental, transport-coupled problem rather than a simple extension of arterial blood gas interpretation. A Stewart framework is particularly useful in CSF because low weak-acid buffering makes $PCO_2$ and SID the dominant determinants of pH in many conditions, while barrier transport provides a mechanistic basis for delayed and pathology-sensitive compensation. Classic clinical observations[3,13] and modern paired CSF–arterial datasets[8,9] converge on the practical point that blood and CSF can move on different trajectories. Mechanistic work supports that these trajectories are actively regulated rather than passive[10,11,16]. At present, CSF Stewart analysis is best regarded as an interpretive framework that strengthens mechanistic reasoning and hypothesis generation. Its transition into a validated bedside tool will require time-resolved human data, standardized measurement, and tighter linkage between physicochemical state variables, barrier transport, and ventilatory-control circuitry.

## References


1. Langer, T., Zanella, A. & Caironi, P. Understanding the role of the cerebrospinal fluid in acid–base disorders. *Intensive Care Med.* **42**, 436–439 (2016). 10.1007/s00134-015-4059-8
2. Marks, V., Goldring, S., Gordon, A. & Vecchione, J. J. Cerebrospinal fluid acid–base relationships in ketoacidosis and lactic acidosis. *J. Appl. Physiol.* **35**, 813–819 (1973). 10.1152/jappl.1973.35.6.813
3. Posner JB, P. F. Spinal-fluid pH and neurologic symptoms in systemic acidosis. *N Engl J Med.* **277**, 605–613 (1967). 10.1056/NEJM196709212771201
4. PD., W. Modeling the effects of proteins on pH in plasma. *J Appl Physiol* **86**, 1421–1427 (1999). 10.1152/jappl.1999.86.4.1421
5. Fencl, V. & Leith, D. E. Stewart's quantitative acid–base analysis: applications in biology and medicine. *Respir. Physiol.* **91**, 1–16 (1993). 10.1016/0034-5687(93)90085-O
6. Magder, S. & Emami, A. Practical approach to physical-chemical acid-base management: Stewart at the bedside. *J. Clin. Invest.* **125**, 2801–2807 (2015). 10.1086/682954
7. Gilfix, B. M., Bique, M. & Magder, S. A physical chemical approach to the analysis of acid–base balance in the clinical setting. *J. Crit. Care* **8**, 187–197 (1993). 10.1016/0883-9441(93)90002-3
8. Zadek, F., Zanella, A., Caironi, P. & Langer, T. Cerebrospinal fluid and arterial acid–base equilibria in healthy pregnant women undergoing elective caesarean section. *Br. J. Anaesth.* **129**, 726–733 (2022). 10.1016/j.bja.2022.07.048
9. Langer, T., Zanella, A. & Caironi, P. Cerebrospinal fluid and arterial acid–base equilibrium of spontaneously breathing patients with aneurismal subarachnoid hemorrhage. *Neurocrit. Care* **37**, 102–110 (2022). 10.1007/s12028-022-01450-1
10. Murphy, V. A. & Johanson, C. E. $Na^+/H^+$ exchange in choroid plexus and CSF in acute metabolic acidosis or alkalosis. *Am. J. Physiol.* **258**, F1528–F1537 (1990). 10.1152/ajprenal.1990.258.6.F1528
11. Christensen, H. L. et al. The choroid plexus sodium–bicarbonate cotransporter NBCe2 regulates mouse cerebrospinal fluid pH. *J. Physiol.* **596**, 4709–4728 (2018). 10.1113/JP275489
12. Davies, P. Cerebrospinal fluid sampling technique and Astrup pH and $PCO_2$ values. *J. Appl. Physiol.* **40**, 596–600 (1976). 10.1152/jappl.1976.40.4.596
13. Ohman, J. L. J. et al. The cerebrospinal fluid in diabetic ketoacidosis. *N. Engl. J. Med.* **284**, 283–290 (1971). 10.1056/NEJM197102112840601
14. Christensen, I. B., Olesen, E. T. B., Damkier, H. H. & Praetorius, J. Cerebrospinal fluid secretion and ion transport in the choroid plexus. *Front. Physiol.* **4**, 304 (2013). 10.3389/fphys.2013.00304
15. Damkier, H. H., Nielsen, S. & Praetorius, J. The murine choroid plexus epithelium expresses the $2Cl^-/H^+$ exchanger ClC-7 and $Na^+/H^+$ exchanger NHE6 in the luminal membrane domain. *Am. J. Physiol. Cell Physiol.* **314**, C246–C258 (2018). 10.1152/ajpcell.00145.2017
16. Johnson DC Frankel HM, K. H. Effect of furosemide on cerebrospinal fluid composition. *Respir Physiol.* **56**, 301–308 (1984). 10.1016/0034-5687(84)90066-5
17. Battisti-Charbonney A Fisher J, D. J. The cerebrovascular response to carbon dioxide in humans. *J Physiol.* **589**, 3039–3048 (2011). 10.1113/jphysiol.2011.206052
18. López-Barneo, J. in *Respiratory Neurobiology: Physiology and Clinical Disorders, Part I* (eds Chen, R. & Guyenet, P. G.) 73–102 (Elsevier, 2022).

19. Mulkey, D. K., Moreira, T. S., Takakura, A. C., Jahanbani, S. & Olsen, M. L. Retrotrapezoid nucleus chemoreception: mechanisms of function and contributions to disordered breathing in disease. *Trends Neurosci.* **48**, 706–720 (2025). 10.1016/j.tins.2025.07.006

20. Sears, M. R., Sargent, J. C., Hutchison, A. A., Halliday, H. L. & Reid, M. M. Respiration and circulation in experimental meningitis. *J. Clin. Invest.* **54**, 92–102 (1974). 10.1172/JCI107740

21. Nattie, E. E. & Tenney, S. M. Effects of potassium depletion on control of breathing in awake rats. *Am. J. Physiol.* **231**, 588–592 (1976). 10.1152/ajplegacy.1976.231.2.588

22. Hughes R, B. M. A simplified bedside approach to acid-base: fluid physiology utilizing classical and physicochemical approaches. *Anaesth Intensive Care Med.* **14**, 445–452 (2013). 10.1016/j.mpaic.2013.07.013

23. Loeschcke, H. H. & Sugioka, K. pH of the cerebrospinal fluid in cisterna magna and pulmonary ventilation (in German). *Pflugers Arch.* **311**, 99–102 (1969). 10.1007/BF00586927

Submitted: 11/04/2025
Accepted: 02/22/2026
Published: 04/10/2026

Research Article

[1]McGill University

Keywords

tyrosol, farnesol, Candida albicans, Pseudomona aeruginosa, biofilm, quorum-sensing, antimicrobial resistance

Email Correspondence

bianka.dusseault@mail.mcgill.ca

https://doi.org/10.26443/msurj.v21i1.371



Dusseault, Bianka[1]

# Effects of *Candida albicans*-derived Farnesol and Tyrosol on Quorum-Sensing Pathways in Pseudomonas aeruginosa Biofilms: Implication for Antimicrobial Resistance

## Abstract

In 2017, amid an antimicrobial crisis, the World Health Organization classified *Pseudomonas aeruginosa* as a priority pathogen for the Research and Development of new antibiotics. *P. aeruginosa*'s ability to form biofilms–structured aggregation of microcolonies embedded in a self-generated matrix–and regulate virulence through quorum-sensing (QS) often results in antibiotic failure, especially in nosocomial settings. As synthetic antibiotics have become increasingly ineffective in eradicating *P. aeruginosa* biofilm, chronic infections persist in cystic fibrosis and burn wound patients. This review investigates *C. albicans*-derived molecules, tyrosol, and farnesol, as potential quorum-sensing inhibitors (QSIs) of *P. aeruginosa* QS circuits (LasI/LasR, RhlI/RhlR, and PQS) in single-species biofilms. These fungal-derived compounds have been shown to have anti-biofilm and antibacterial activity by disrupting *P. aeruginosa* QS pathways and modulating virulence factor expression. In combinatorial therapeutics, tyrosol and farnesol have been demonstrated to facilitate certain antibiotic activity, suggesting potential for clinical implementation. Unlike traditional antibiotics, *C. albicans* QSIs have co-evolved with *P. aeruginosa* and developed inhibition mechanisms that minimize the selective pressures driving antimicrobial resistance. By targeting cell-to-cell communication rather than bacterial growth, tyrosol and farnesol offer a propitious avenue for *P. aeruginosa* biofilms-associated infections. However, as this remains an emerging field of research in the context of *P. aeruginosa*-driven infections, further research is needed to determine clinical plausibility, especially in *in vivo* models, to understand specific mechanisms of action, dosage optimization and potential undesirable off-target interactions.

## Introduction

In hospital settings, 80% of microbial infections are associated with biofilms, which are 1000-fold more tolerant to antibiotic exposure than their planktonic counterparts[1, 2, 3, 4]. Amid an antimicrobial resistance crisis, *Pseudomonas aeruginosa*, a gram-negative bacterium, has become increasingly challenging to treat with conventional therapeutics, often leading to antibiotic failure[3]. In nosocomial settings, its ability to form biofilms results in chronic infections in patients with cystic fibrosis, burn wounds, and medical devices (e.g., urinary catheters and implants)[1, 4, 5, 6, 7]. A biofilm is a complex, compact community of bacteria embedded in a self-produced extracellular matrix (EPS) made of polysaccharides, extracellular DNA, proteins, and lipids[1, 3, 4]. Constituting 90% of the biofilm mass, the matrix acts as a protective barrier for microcolonies against environmental stresses, allowing the adaptation of *P. aeruginosa* to extreme niches[3]. Internally, EPS facilitates nutrient flow, antibiotic tolerance, and cell-to-cell communication[1]. Historically, planktonic bacteria have been studied as primitive, solitary organisms lacking the social and behavioral mechanisms observed in eukaryotes. However, in recent decades, research has redefined them as complex social communities that interact, coordinate, and communicate through systems known as quorum-sensing (QS)[8, 9]. QS systems are concentration-dependent mechanisms that operate through small molecules called autoinducers to regulate gene expression, antibiotic production, virulence factors, and biofilm formation[1, 2, 8]. In *P. aeruginosa*, three distinct QS systems, LasI/LasR, RhlI/RhlR, and PQS, are utilized by bacteria to evade the host immune system and produce public goods (e.g., virulence factors) that increase antibiotic tolerance[3, 4, 9].

As the failure of synthetic antibiotics has become increasingly common, researchers are turning to natural compounds as promising alternatives. Due to the coevolution of bacteria and fungi in their ecological niche, fungi like *C. albicans* have been able to produce quorum-sensing molecules that may interfere with *P. aeruginosa* quorum-sensing pathways[10]. Among these, tyrosol and farnesol have emerged as potential quorum-sensing inhibitors (QSIs) with implications for antibiotic stewardship. While tyrosol and farnesol have been extensively investigated in the context of *C. albicans* biofilm formation, little research has studied their effects on other microorganisms and their potential therapeutic applications. This review demonstrates that tyrosol and farnesol have significant clinical potential as adjuncts to conventional antibiotics. By modulating quorum-sensing pathways and virulence factors expression rather than bacterial growth, they enhance antibiotic susceptibility while minimizing selective pressures for resistance development. As this avenue remains largely unexplored, we will critically analyze current limitations and propose future research directions.

2. Methodology

A literature search was conducted using Pubmed, and Scopus. Search terms included "farnesol", "tyrosol", "quorum-sensing", "quorum-quenching", "*Pseudomonas aeruginosas*" and "*Candida albicans*", among others. Articles published in the last 20 years were considered. Titles were screened for relevance and studies focused on investigating the effect of farnesol and tyrosol on *P. aeruginosa* biofilm formation, virulence factors, and combinatorial therapeutics were included. Papers published in another language than English were excluded. A total of 27 papers were included, with 16

primary studies, 10 reviews, and 1 book chapter.

## 3. Quorum-Sensing Systems in *P. aeruginosa*

Quorum-sensing is a density-dependent mechanism that uses chemical signals to enable the coordination of social behaviors within bacterial communities. The *P. aeruginosa* QS system contains three interdependent circuits–LasI/LasR, RhlI/RhIR, and PQS–with LasI and RhlI approximately regulating 10% of *P. aeruginosa* genome[3, 4, 9, 11]. The LasI/LasR system typically regulates the RhlI/RhIR, which then influences PQS production, while PQS is an intermediary that modulates both AHL-mediated systems[12]. The LasI/LasR and RhlI/RhIR use acyl-homoserine lactone signals, whereas the PQS system operates with quinolone signaling[3, 4, 9]. Relevant to this review is the contribution of QS to biofilm formation in *P. aeruginosa*, including the expression of genes encoding for proteases, toxins, and virulence factors at bacterial quorum size. QS-regulated products can influence group behaviors like nutrient acquisition, microbial competition, antibiotic tolerance, cytotoxicity, and pathogenicity[4]. This enables the biofilm to persist in a sessile stage, exhibiting a 1000-fold more tolerance to current antibiotic treatment than planktonic bacteria[2, 3, 4].

### 3.1 AHLs System

The AHL-mediated quorum-sensing systems in *P. aeruginosa* consist of two circuits: LasI/LasR and RhlI/RhIR, both regulated by LuxI-like synthases, LasI and RhlI, respectively[3, 4, 9]. LasI synthase produces the signal molecule 3OC12-HSL (N-3-oxo-dodecanoyl homoserine lactone), which binds to the transcriptional regulator LasR at quorum-sensing threshold levels. Similarly, RhlI synthase produces C4-HSL (N-butyryl homoserine lactone), which binds to RhIR[34, 11]. These interactions result in the production of LuxR-like transcription factors, functioning mainly as activators of QS-regulated genes[9]. The LasI/LasR and RhlI/RhIR systems regulate the expression of virulence factors in distinct and independent ways, including protease elastase (*lasB* gene) and rhamnolipids, respectively[3]. Rhamnolipids, a biosurfactant, preserves mature biofilm architecture by maintaining pores and channels between microcolonies, facilitating nutrient and liquid flow within the matrix[9]. Additionally, they contribute to immune evasion by inducing lytic necrosis in polymorphonuclear neutrophils (PMNs) and macrophages[3, 4]. It is worth noting that QS-regulated genes can rely on either LasR or RhlR, while others, like *lasB*, rely on both circuits[9].

### 3.2 PQS System

The PQS system, the most recently discovered, utilizes two signaling molecules: high-affinity PQS (2-heptyl-3-hydroxy-4-quinolone) and low-affinity HHQ (2-heptyl-4-hydroxyquinoline)[3, 4]. These chemical signals activate the transcriptional regulator *PqsR*, leading to the expression of the *pgsABCDE* operon. PQS operates through a self-reinforcing feedback loop, similar to the LasR-LasI AHL system[4, 9]. Acting as an intermediary system in the quorum-sensing hierarchy of *P. aeruginosa*, PQS is activated by LasR but inhibited by RhlR. Work has also shown that the gene product of *pqsE*, while not directly involved in PQS, can regulate RhlR activity, but its mechanistic inhibition remains unclear[9]. The QS-regulated products of the PQS system are lectin A, hydrogen cyanide, and pyocyanin[3]. Pyocyanin is of particular interest as it is responsible for approximately 95% of *P. aeruginosa* antimicrobial properties and promotes biofilm formation[13].

## 4. Effect of Tyrosol and Farnesol in *P. aeruginosa*

Farnesol and tyrosol, *C. albicans*-derived compounds, are potential QSIs that can disrupt quorum-sensing pathways in *P. aeruginosa*. In *C. albicans*, they have been shown to work in tandem during biofilm formation and morphological growth; tyrosol promotes the transition from yeast to hyphae while farnesol inhibits this shift at high cell densities[2, 10, 14, 15]. This dynamic regulates *C. albicans* pathogenicity and biofilm structure, shaping its response to different ecological niches and natural competitors like *P. aeruginosa*[10]. In polymicrobial infections, *C. albicans* and *P. aeruginosa* not only co-exist but also engage in bidirectional communication. Research has shown that *P. aeruginosa* can impede fungal growth through 3OC12-HSL, a QS signaling molecule, while *C. albicans* can disrupt PQS signal and associated virulence factor (e.g. pyocyanin) through farnesol[11, 12, 14, 16, 17]. In *Pseudomonas aeruginosa*-driven chronic infections, tyrosol and farnesol are promising therapeutics, either in isolation or combined with antimicrobials. For example, in PAO1 biofilms, farnesol emulsion of 3mg/ml not only disrupted biofilm formation, but it also preserved lung cell viability more effectively than the antibiotic ciprofloxacin[18]

### 4.1 Disruption of Quorum-Sensing Pathways

QPCR analysis of non-mucoid (NCTC 10,662), mucoid (PAO1), and heavily mucoid (RBHi) strains of *Pseudomonas aeruginosa* demonstrates that *C. albicans*-derived compounds, farnesol, and tyrosol, decrease the expression of LasI and *RhlL*[10, 19]. In turn, this downregulates the production of AHLs signal molecules, 3OC12-HSL and C4-HSL. Furthermore, at α=0.05, farnesol appears to have a stronger effect on the inhibiting of LasI compared to tyrosol[10]. This suggests that farnesol may be more effective in disrupting the LasI/LasR. However, farnesol shows to up-regulate LasR activity in both strains NCTC 10,622, and PAO1, whereas tyrosol only increased LasR activity in the latter[10]. This finding highlights the strain specificity of quorum sensing modulators. Another study demonstrates that tyrosol decreased C4-HSL and 3OC12-HSL by 72%, and 75% at 1/4x MIC, respectively in *P. aeruginosa*[19]. The interactions of farnesol and tyrosol with the QS-system in *P. aeruginosa* warrants further investigation, as the mechanisms of actions under those compounds operating in various bacterial strains remain unclear. While these compounds can be beneficial, they can have inadvertent effects in specific contexts, such as enhancing biofilm persistence or virulence, which makes their use in treatment strategies rather unpredictable.

### 4.2 Regulation of virulence factors and growth

In recent years, it has been shown that compounds like farnesol exhibit anti-biofilm properties, whereas tyrosol possess antibacterial activity, although some studies suggest that farnesol does exhibit antibacterial activity as well[8, 10, 19, 20, 21, 22, 23, 24]. These properties, respectively, halt or slow the progression of the biofilm and limit the growth and viability of bacteria. Additional research has indicated that farnesol disrupts the PQS system of *P. aeruginosa* by inhibiting 2-heptyl-3-hydroxy-4-quinolone[2] and represses *pqsA* transcription, first gene of the PQS operon, in a dose-dependent manner[14]. This leads to a reduction in the production of pyocyanin, a PQS-regulated virulence factor, that decreases bacteria's swarming motility by lowering haemolysin production[2, 16]. It is worth noting that, as well as inhibiting protease production and impacting swarming motility, tyrosol inhibits heamolysin production to a greater degree than farnesol[2, 8, 19]. Specifically, testing on 20 *P. aeruginosa* isolates at α=0.001, 1.2 μM tyrosol significantly decreases both heamolysin and protease production whereas 10 μM farnesol only affects heamolysin production[8]. Furthermore, tyrosol at 10 μM inhibited >50% of *P. aeruginosa* growth after a 16h cultivation, whereas farnesol concentrations of 200 μM inhibited >30% growth[8]. It is important to note that anti-resistance advantages are solely present when QSIs are used at their sub-inhibitory concentrations, which disrupt QS signaling pathways without affecting cell viability. Concentrations lower than 1.2 μM tyrosol and 10 μM farnesol have little to no effect on *P. aeruginosa* growth[8]. However, farnesol concentrations of 25 μM effectively disrupt the PQS system[14]. Further research by Cugini *et al.* demonstrates that farnesol can stimulate PQS production in *P. aeruginosa* LasR-defective mutants,

**Table 1.** Summarizing key studies evaluating the effects of farnesol and tyrosol on P. aeruginosa strains

| Strain Type | Experimental Model | Treatment | Key Findings | Study |
|---|---|---|---|---|
| PAO1 | In vitro single-species biofilms of *P. aeruginosa* | Farnesol/ tyrosol | Farnesol showed stronger antibiofilm activity | (Hacioglu *et al.*, 2024) |
| | In vitro dual-species biofilms (*C. albicans* + *P. aeruginosa*) | Farnesol/ tyrosol | Both reduced biofilm counts, especially farnesol | |
| Clinical isolates | *In vitro* | Farnesol/ tyrosol | Growth inhibition and reduced haemolysin and protease production at certain concentrations | (Abdel-Rhman *et al.*, 2015) |
| NCTC 10,662<br>PAO1<br>RBHi | *In vitro* | Tyrosol and farnesol with furanone treatment | Up-regulation of LasR with farnesol<br>Up-regulation of LasR with tyrosol and farnesol<br>Up-regulation of RhlR receptor with tyrosol<br>Shared results included:<br>• Significant reduction in LasI protein with farnesol > tyrosol;<br>• Reductions in mRNA expression for RhlI protein;<br>• Down-regulation of toxA, aprA, rhlAB and LasB | (Kalgudi *et al.*, 2022) |
| PA14 | *In vitro* | Farnesol | Reduction in pyocyanin production and dose-dependent reduction in PQS production by inhibition of *pqsA* | (Cugini, 2007) |
| | *In vitro C. albicans* and *P. aeruginosa* co-cultures | Farnesol | Reduction in pyocyanin and PQS production | |
| PAO1 | *In vitro* | Farnesol | Reduction in pyocyanin production | |
| ΔLasR | *In vitro* | Farnesol | Stimulation of PQS production | (Cugini *et al.*, 2010) |

which are frequently observed in cystic fibrosis infections[25]. This finding further underscores that QSIs may exhibit strain-specific effects which requires extensive investigation when developing targeted therapeutic interventions. Another study demonstrates that farnesol and tyrosol can down-regulate virulence-related genes that code for the proteins toxA, aprA, LasB, and rhlAB, which are key to the production of rhamnolipids[10]. Findings are summarized in Table 1.

## 5. Implications for Antimicrobial Resistance

In the past decades, *P. aeruginosa* has become increasingly resistant to many classes of antibiotics due to its high versatility and ability to form persistent biofilms[3]. In nosocomial infections, broad spectrum antibiotics result in antibiotic failure, often giving rise to chronic diseases. This can lead to more invasive procedures such as surgical debridement, a standard method for treating chronic wounds[5, 6, 7, 20]. However, natural QS-agents like farnesol and tyrosol offer a promising avenue that can co-evolve with *P. aeruginosa* in their ecological niche[10]. Evidence suggests that *C. albicans* and *P. aeruginosa* have mutually antagonistic effects that ultimately increase mutability rates in both species, favoring constant evolution[16, 17] .

Contrary to synthetic antibiotics, compounds like farnesol and tyrosol can naturally develop counter-adaptations to *P. aeruginosa* resistance mechanisms, and if harnessed, can provide novel therapeutics. If combined with other pharmacological products, QS-agents, including farnesol and tyrosol, might enhance antibiotic susceptibility by disrupting quorum-sensing pathways and regulating virulence factors. This facilitates the penetration of antibiotics through the biofilm's extracellular matrix. It is widely believed that QS-agents, farnesol and tyrosol, do not interfere with the growth of the biofilm when used at their sub-inhibitory concentrations. This minimizes the selective pressures for resistance[14], slowing down the emergence of multi-drug-resistant strains of *P. aeruginosa*[4]. While QS-agents present a propitious strategy, considerations concerning their long-term efficacy, potential bacterial interactions and clinical implementation warrant further investigation (discussed in a later section).

### 5.1 Combination therapeutics

Combination therapeutics involve the use of several drugs that synergistically interact to target different biological pathways or mechanisms. By combining natural compounds such as farnesol and tyrosol with antibiotics, we can potentially minimize the occurrence of invasive procedures while mitigating the selective pressures of antimicrobial resistance on bacterial strains. Additionally, natural compounds can have synergistic effects with conventional antibiotics, which can mitigate the dosage of antibiotics, preventing their overuse in medical settings. Multiple studies have demonstrated that those compounds enhance antibiotic susceptibility in organisms such as *E. coli*, *S. aureus* and *P. aeruginosa*[12, 21, 26].

Hacioglu *et al.* in-vitro study demonstrates that the combination of farnesol with antibiotics, particularly colistin (200 mg/L), produced a statistically significant effect at α=0.05, whereas tyrosol combinations did not yield significant results in single-species *P. aeruginosa* biofilms[2]. However, Abdel-Rhman S.H. *et al.* reported that tyrosol and farnesol, when used at their subinhibitory concentrations, did not have a significant effect on the antibacterial activity of ceftriaxone, ciprofloxacin, and gentamicin on *P. aeruginosa* at concentrations of 64000-0.12, 4000-0.004, and 2000-0.002 μg/mL respectively[8]. Taking it a step further, Bandara *et al.* demonstrated that liposomal ciprofloxacin combined with farnesol disrupted extensively the biofilm's structure and lowered metabolic activity. This resulted in a significant decrease in viability[27]. Using the *P. aeruginosa* TL2314 strain in *G. mellonella*, Han *et al.* (2023) shows in-vitro and *in vivo* synergistic antibacterial effects between colistin and farnesol. While this review focuses on single-species *P. aeruginosa* biofilms, an *in vivo* study using *C. elegans* model shows that the combination of tyrosol or farnesol with colistin is statistically significant, at α=0.05 in dual-species *C. albicans*-*P. aeruginosa* biofilms[2]. Specifically, the combination of farnesol and colistin was most effective, which supports in vitro results in single-species *P. aeruginosa* biofilms[2]. Additionally, combinations of farnesol and tyrosol with fluconazole (10 mg/L) increased *C. elegans* viability and proved to be more effective than isolated QQMs[2]. To date, few *in vivo* studies have tested farnesol and tyrosol with antibiotics in single-species biofilms. Nonetheless, further investigation would be relevant based on the *in vivo* findings in dual-species biofilms. Along with the previous studies, this suggests that

**Table 2.** Summarizing key studies evaluating the effect of combinatorial therapeutics on P. aeruginosa strains

| Strain Type | Experimental Model | Treatment | Key Outcomes | Study |
|---|---|---|---|---|
| PAO1 | *In vitro* single-species biofilms | Tyrosol/ Farnesol; Fluconazole/ amphotericin/ caspofungin/ aztreonam/ colistin/ tobramycin | Farnesol-colistin combination is most effective; all antibiotic and farnesol are statistically significant | (Hacioglu *et al.*, 2024) |
| | *In vitro* dual-species biofilms (*C. albicans* + *P. aeruginosa*) | Tyrosol/ Farnesol; Fluconazole/ amphotericin/ caspofungin/ aztreonam/ colistin/ tobramycin | Farnesol-colistin combination is most effective; all combinations are statistically significant | |
| | *In vivo C. elegans* models (*C. albicans* + *P. aeruginosa*) | Tyrosol/ Farnesol; Colistin/ Fluconazole | Increased survival for any combinations; Farnesol-colistin combination was more effective than tyrosol combinations against *P. aeruginosa* | |
| Clinical isolates | *In vitro* | Farnesol/ Tyrosol; Ceftriaxone/ Ciprofloxacin/ Gentamicin | No significant effect on the antibacterial activity of tested antibiotics when used at their sub-inhibitory concentrations | (Abdel-Rhman *et al.*, 2015) |
| PAO1 | *In vitro* | Farnesol; Liposomal ciprofloxacin | Biofilm disruption and lower live/dead cell ratio | (Bandara *et al.*, 2016) |
| TL2314 | *In vitro* and *in vivo G.mellonella* models | Farnesol; colistin | Synergistic antibacterial effect between farnesol and colistin | (Han *et al.*, 2023) |

tyrosol and farnesol may act synergistically with the tested antibiotics, enabling reduced antibiotic doses without promoting resistance or tolerance in *P. aeruginosa* strains at their specific concentrations[2, 8, 27].

For clinical use, it will be crucial to use effective detection methods of early-stages biofilms to ensure appropriate treatment as the effectiveness of combinatorial therapeutics can depend on the maturation stages of biofilms. Therefore, treatments applied during the early stages of biofilm formation are usually more effective than those applied to mature biofilms. Immature biofilms usually exhibit greater susceptibility to antibiotics, whereas mature biofilms tend to be more resistant due to structural and physiological changes, such as the development of the extracellular matrix[5, 6]. Findings are summarized in Table 2.

## 6. Future Direction and Research Gaps

The research of natural compounds as potential modulators of the quorum-sensing system is a growing and active field of research. Yet, there are many research gaps that need to be addressed, ultimately informing the research processes.

Since very few studies have been conducted on *in vivo* models, there is a lack of knowledge on the biological plausibility of farnesol and tyrosol in a medical setting. *In vivo* studies will be useful in examining fully the phenotypic and genotypic complexity of biofilm growth in diverse micro-environment[7]. Therefore, they will be key in determining the mechanisms of actions of quorum-sensing inhibitors, dose-response, and potential off-target interactions, all of which remain unclear. Although not in the scope of this mini-review, it is also worth noting that future research on dual-species biofilms would be relevant to understanding potential undesirable off-target synergistic effects at play between *C. albicans* and *P. aeruginosa*. There is also a need to assess the stability and bioavailability of farnesol and tyrosol to determine if pharmaceutical production is plausible and possible at the global stage. Furthermore, there is the possibility that *P. aeruginosa* strains develop resistance to QS-interfering agents themselves. However, this is a challenge pertinent to any new and current compounds, synthetic or natural. The development of resistance is an evolutionary process that cannot be evaded but slowed down by a smart distribution and usage of antimicrobials as well as diversification of treatments.

As discussed in a previous section, combination therapeutics may be used adjunctively with standard care to develop targeted interventions that could minimize invasive procedures when infections are detected at an early stage[5, 7]. However, within the scope of this review, *in vivo* and *in-vitro* studies on the effect of tyrosol and farnesol used conjointly with antibiotics in single-species *P. aeruginosa* biofilms were sparse. To investigate the role of QSIs in antimicrobial stewardship, we need to explore the potential for combinatorial therapies and determine the most effective combinations specific to *P. aeruginosa* strains[2]. Kalgudi *et al.* show that, mucoid, non-mucoid, and heavily mucoid *P. aeruginosa* strains responded differently to various combinations. This highlights the need for testing the strain-specific effect of farnesol and tyrosol as *Pseudomonas aeruginosa* is an extremely versatile and adaptive organism that can express different phenotypes.

In retrospect, natural compounds like tyrosol and farnesol have demonstrated anti-biofilm and anti-bacterial activity in single-species *P. aeruginosa* biofilms, primarily through the quorum-sensing dysregulation and the regulation of the expression of virulence factors. As antimicrobial resistance continues to rise, quorum-sensing inhibitors could open a new niche within antibiotic stewardship, offering dynamic alternative treatments.

## References


1. Sharma, D., L., M. & A.U., K. Antibiotics versus biofilm: an emerging battleground in microbial communities. *Antimicrob. Resist. Infect. Control* **8** (2019). 10.1186/s13756-019-0533-3
2. Hacioglu, M., Yilmaz, F., Yetke, H. & Haciosmanoglu-Aldogan, E. Synergistic effects of quorum-sensing molecules and antimicrobials against candida albicans and pseudomonas aeruginosa biofilms: in vitro and in vivo studies. *J. Antimicrob. Chemother.* **79** (2024). 10.1093/jac/dkae293
3. Thi, M., Wibowo, D. & Rehm, B. Pseudomonas aeruginosa biofilms. *Int. J. Mol. Sci.* **21** (2020). 10.3390/ijms21228671
4. Jakobsen, T., Bjarnsholt, T., Jensen, P., Givskov, M. & H□iby, N. Targeting quorum sensing in pseudomonas aeruginosa biofilms: Current and emerging inhibitors. *Future Microbiol.* **8** (2013). 10.2217/fmb.13.57
5. Shen, A., Taha, M., Ghannoum, M. & Tyring, S. Biofilms and Chronic Wounds: Pathogenesis and Treatment Options. *J. Clin. Med.* **14,** 7784 (2025). 10.3390/jcm14217784
6. Shultz, G. et al. Consensus guidelines for the identification and treatment of biofilms in nonhealing wounds. *Wound Repair Regen.* **25,** 744–757 (2017). 10.1111/wrr.12590

7. Weigelt, M., McNamara, S., Sanchez, D., Hirt, P. & Kirsner, R. Evidence-Based Review of Antibiofilm Agents for Wound Care. *Adv. Wound Care (New Rochelle)* **10,** 13–23 (2021). 10.1089/wound.2020.1193

8. Abdel-Rhman, S., Mostafa El-Mahdy, A. & El-Mowafy, M. Effect of tyrosol and farnesol on virulence and antibiotic resistance of clinical isolates of pseudomonas aeruginosa. *BioMed. Res. Int.* **1–7** (2015). 10.1155/2015/456463

9. Miranda, S., Asfahl, K., Dandekar, A. & Greenberg, E. Pseudomonas aeruginosa quorum sensing. *Adv. Exp. Med. Biol.* **1386,** 95–115 (2022). 10.1007/978-3-031-08491-1_4

10. Kalgudi, R., Tamimi, R., Kyazze, G. & Keshavarz, T. Effect of quorum quenchers on virulence factors production and quorum sensing signalling pathway of non-mucoid, mucoid, and heavily mucoid pseudomonas aeruginosa. *World J. Microbiol. Biotechnol* **38** (2022). 10.1007/s11274-022-03339-9

11. Williams, P. & Cámara, M. Quorum sensing and environmental adaptation in Pseudomonas aeruginosa: a tale of regulatory networks and multifunctional signal molecules. *Curr. Opin. Microbiol.* **12,** 182–191 (2009). 10.1016/j.mib.2009.01.005

12. Sordi, L. & Mühlschlegel, F. A. Quorum sensing and fungal-bacterial interactions in Candida albicans: a communicative network regulating microbial coexistence and virulence. *FEMS Microbiol. Ecol.* **9,** 990–999 (2009). 10.1111/j.1567-1364.2009.00573.x

13. Mudaliar, S. & Prasad, A. A biomedical perspective of pyocyanin from Pseudomonas aeruginosa: its applications and challenges. *World J. Microbiol. Biotechnol.* **40** (2024). 10.1007/s11274-024-03889-0

14. Cugini, C. et al. Farnesol, a common sesquiterpene, inhibits PQS production in Pseudomonas aeruginosa. *Mol. Microbiol.* **65,** 896–906 (2007). 10.1111/j.1365-2958.2007.05840.x

15. Alem, M., Oteef, M., Flowers, H. & Douglas, J. Production of Tyrosol by Candida albicans Biofilms and Its Role in Quorum Sensing and Biofilm Development. *Eukaryot. Cell* **5,** 1770–1779 (2006). 10.1128/EC.00219-06

16. McAlester, G., O'Gara, F. & Morrissay, J. Signal-mediated interactions between Pseudomonas aeruginosa and Candida albicans. *J. Med. Microbiol.* **57,** 563–569 (2008). 10.1099/jmm.0.47705-0

17. Trejo-Hernández, A., Andrade-Domínguez, A., Hernández, M. & Encarnación, S. Interspecies competition triggers virulence and mutability in Candida albicans–Pseudomonas aeruginosa mixed biofilms. *ISME J.* **8,** 1974–1988 (2014). 10.1038/ismej.2014.53

18. Tan, L., Sutton, K., Murphy, S. & Levi, N. Farnesol emulsion for elimination of Pseudomonas aeruginosa biofilm in a 3D airway model for cystic fibrosis. *Biofilm* **10** (2025). 10.1016/j.bioflm.2025.100317

19. Abdel-Rhman, S., Rizk, D. & Abdelmegeed, E. Effect of Sub-Minimum Inhibitory Concentrations of Tyrosol and EDTA on Quorum Sensing and Virulence of Pseudomonas aeruginosa. *Infect. Drug Resist.* **13,** 3501–3511 (2020). 10.2147/IDR.S264805

20. Bayer, I. Fungal quorum sensing molecules as potential drugs in the treatment of chronic wounds and their delivery. *Expert Opin. Drug Deliv.* **22,** 277–296 (2025). 10.1080/17425247.2025.2452303

21. Jabra-Rizk, M., Meiller, T., James, C. & Shirtliff, M. Effect of Farnesol on Staphylococcus aureus Biofilm Formation and Antimicrobial Susceptibility. *Antimicrob. Agents Chemother.* **50,** 1463–1469 (2006). 10.1128/AAC.50.4.1463-1469.2006

22. Sun, A. et al. Tyrosol from Marine Fungi, a novel Quorum-Sensing inhibitor against Chromobacterium violaceum and Pseudomonas aeruginosa. *Bioorg. Chem.* **91** (2019). 10.1016/j.bioorg.2019.103140

23. Silva, B., Miguel, R., Miletti, L. & Quadros, R. Inhibitory effects of EDTA, farnesol and xylitol on biofilms produced by Staphylococcus aureus and Pseudomonas aeruginosa. *Microbe* **7** (2025). 10.1016/j.microb.2025.100321

24. Zhang, Y. et al. Synergy with farnesol rejuvenates colistin activity against Colistin-resistant Gram-negative bacteria in vitro and in vivo. *Int. J. Antimicrob. Agents* **62** (2023). 10.1016/j.ijantimicag.2023.106899

25. Cugini, C., Morales, D. & Hogan, D. Candida albicans-produced farnesol stimulates Pseudomonas quinolone signal production in LasR-defective Pseudomonas aeruginosa strains. *Microbiology* **156,** 3096–3107 (2010). 10.1099/mic.0.037911-0

26. Choi, H.-Y. & Kim, W.-G. Tyrosol blocks E.Coli anaerobic biofilm formation via YbfA and FNR to increase antibiotic susceptibility. *Nat. Commun.* **15** (2024). 10.1038/s41467-024-50116-3

27. Bandara, H. et al. Incorporation of Farnesol Significantly Increases the Efficacy of Liposomal Ciprofloxacin against Pseudomonas aeruginosa Biofilms in Vitro. *Mol. Pharm.* **13,** 2760–2770 (2016). 10.1021/acs.molpharmaceut.6b00360