
MSURJ
MCGILL SCIENCE UNDERGRADUATE RESEARCH JOURNAL
VOLUME 20 | ISSUE 1 | APRIL 2025

ON THE COVER

What would it be like to see into the brain? To map its pulses of electricity, traveling across networks upon networks of branching, interwoven neurons?

MSURJ's 20th anniversary cover explores the possibilities of visualizing brain activity through electroencephalography (EEG), illustrating its potential applications to assessing consciousness recovery in sedated pediatric epilepsy patients. Inspired by author Marlo Naish et. al.'s innovative work, found on pages 1-8, this cover artwork offers a colorful and fantastical interpretation of the ebbs and flows of chaotic, seizure-state brain waves.

Typically, "windows" into the brain look more like sticky electrodes with wires that transmit electrical signals from the head to a computer. However, artist Emma Lowry interprets them as eyes peering into the world of layered, textured signs of life, which are stronger in some areas–indicated in bright red–and weaker in others, shown in cool blue tones. Similar to a topographical map, she localizes the highest complexity EEG indicators in the warmest shades, reflecting their intricate potential as predictors of the restoration of consciousness. Created with a combination of acrylic paint markers and digital art tools, this drawing is a whimsical portrait of the artist's imagining of the hectic and resilient mind of the child.

Emma Lowry is a fourth-year McGill student studying Gender and Feminist Studies and Biochemistry. She is passionate about creating accessible scientific illustrations and colorful, playful anatomical artwork.

MCGILL SCIENCE
UNDERGRADUATE
RESEARCH JOURNAL

BOARD 2024-2025

805 Sherbrooke St. West, Room 1B22
Montreal, Quebec, H3A 2K6 Canada
Phone: (514) 398-6979
Fax: (514) 398-6766
Email: msurj.sus@mail.mcgill.ca
Website: https://msurjonline.mcgill.ca/about/editorialTeam

TABLE OF CONTENTS

FOREWORD

Dear Reader,

As we celebrate the twentieth anniversary of the *McGill Science Undergraduate Research Journal*, we are struck by one force that has remained constant throughout two decades of scientific inquiry: curiosity. This sincere desire to better understand the world drives students towards scientific exploration. Since our founding in 2005, we've provided a platform where undergraduate researchers can transform their work into peer-reviewed contributions to scientific knowledge. The same spark of inspiration that led to our founding continues to drive the work of each undergraduate researcher published in these pages.

This milestone volume features eight exceptional articles spanning diverse scientific disciplines—from mathematics and complex systems biology to conservation ecology and neuroscience. We are particularly proud to showcase innovative research on neural complexity markers in pediatric epilepsy, de Sitter vacua in string theory, and approaches to climate adaptation in Québec's dairy industry. These works exemplify how today's undergraduate researchers are tackling both fundamental scientific questions and pressing real-world challenges with remarkable depth and insight.

With twenty volumes behind us, we recognize that our journal serves not just as a publication platform, but as a space where students engage with science beyond inquiry and data collection. We have been fortunate to work with such a dedicated and talented Editorial Board, whose commitment to scientific rigor has maintained the high quality of our journal. Through coordinating peer review, our editors ensure each paper meets the high standards of scientific publication. While upholding these editorial standards, our team has also collaborated on new seminars, workshops, and outreach events throughout the year that have advanced our support of undergraduate research in an unprecedented manner. This collaborative and dynamic spirit of scientific excellence continues to be the foundation of our publication.

We present to you *Volume 20* with immense pride in our past and excitement for the future of undergraduate scientific inquiry.

On behalf of our entire Editorial Board, thank you.

Mina Mahdi and Armance Volta
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 Medicine and Health Sciences
Centre for Structural Biology Research
Department of Physiology
Department of Earth and Planetary Sciences
Department of Mathematics and Statistics
Department of Physics
Department of Atmospheric and Oceanic Sciences
Department of Geography

Finally, we thank the peer reviewers, including Dr. Joydeep Chakravarty, Dr. Khanh Huy Bui, Dr. Jean-François Trempe, Olivia Zhu, Dr. Laurent Sansregret, Luke McCaffrey, Dr. Zhengchen Cai, and the anonymous peer reviewers. Their careful revision of these articles is critical to the success of our 20th volume.

Submitted: 01/16/2025
Accepted: 03/06/2025
Published: 03/31/2025

Research Article

[1]Faculty of Science, McGill University, Montréal, QC, Canada

[2]Integrated Program in Neuroscience, McGill University, Montréal, QC, Canada

[3]School of Physical and Occupational Therapy, McGill University, Montréal, QC, Canada

Keywords

Epilepsy, Consciousness, Pediatrics, Electroencephalography, Prognosis

Email Correspondence

marlo.naish@mail.mcgill.ca

https://doi.org/10.26443/msurj.v20i1.288

Marlo Naish[1], Derek Newman[2] and Stefanie Blain-Moraes[3]

Neural Complexity and Prognosis: Predicting Recovery in Pediatric Epilepsy Using EEG Markers

Abstract

Predicting patient functional outcomes is an indispensable part of clinical care in the Pediatric Intensive Care Unit (PICU), especially for children with epilepsy, a prominent neurological emergency. Electroencephalography (EEG) is a dynamic tool for assessing brain activity, with brain complexity and spectral power features emerging as predictors of consciousness recovery. We investigated whether patients' EEG activity under anesthesia could predict their recovery, using data from 12 pediatric epilepsy patients (mean age: 11.0±2.2 years). Neural complexity, the intricacy of connectivity between brain regions, is heavily implicated in a patient's capacity for consciousness. We hypothesize that neural complexity will be a stronger predictor of patient outcomes than spectral power and that higher complexity will be associated with better outcomes. EEG features were analyzed during sedated, baseline (non-sedated), and difference states. Recovery was assessed three months post-injury using the Glasgow Outcome Scale-Extended (GOS-E). The predictive performance of significant EEG markers was evaluated using logistic regression with leave-one-out cross-validation and permutation testing. Baseline EEG features showed minimal prognostic power, whereas sedation and difference states yielded high prognostic accuracy. In the sedated state, the complexity features rate entropy and Lopez-Ruiz-Mancini-Calbet Complexity (HC-LMC) predicted recovery, separating good and poor outcomes with 100% accuracy. These findings demonstrate that EEG markers of complexity can predict the recovery of consciousness in pediatric epilepsy patients under anesthesia. Therefore, EEG analysis could be an accessible, accurate, and powerful prognostic tool in clinical settings. Future research should explore these results in larger samples to validate the findings that rate entropy and HC-LMC are predictive of recovery. Further, these features should be studied in patients of different etiologies to analyze their potential as generalizable markers of consciousness.

Introduction

An unresponsive, brain-injured patient's capacity for consciousness is an integral component of their prognosis. Accurate, efficient prognostication allows clinicians to develop informed care strategies and recovery plans. The need for a powerful tool to assess the recovery of consciousness is particularly evident in the Pediatric Intensive Care Unit (PICU), which provides care to children with critical conditions that often lead to states of unconsciousness. Epilepsy, a common reason for admission to the PICU, is a condition characterized by frequent seizures caused by abnormal electrical activity in the brain. Seizures cause transient or prolonged disruptions in consciousness and are the most prevalent childhood neurological emergency[1,2].

Neural complexity, the complexity of the brain's connectivity, is a key framework for understanding consciousness. Early studies suggest that high complexity necessarily supports awareness[3]. This idea has been corroborated by recent studies showing that brain signal complexity correlates with consciousness levels: in both electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) analysis, complexity decreased with anesthesia-induced unconsciousness[4,5].

Anesthesia is commonly administered to PICU patients for sedation and pain management, and neural changes with anesthesia-induced unconsciousness can be monitored with EEG (a non-invasive measure of the brain's electrical activity)[6,7]. These changes arise from the brain's reconfiguration with anesthesia, adapting to a new physiological environment by shifting connectivity. In the injured adult brain, the extent of this reconfiguration reflects the brain's capacity to recover consciousness[8]. Therefore, assessing the brain's EEG signature under anesthesia is a promising method of predicting a patient's recovery.

EEG is often used to assess the neural status of patients with epilepsy[1,9]. Disruptions of consciousness during epileptic seizures are correlated with a loss of EEG complexity[10]. This finding supports the conclusions from anesthesia research: a loss of consciousness is associated with reduced neural complexity. As shown with anesthesia-induced unconsciousness, a behaviourally unresponsive patient's EEG activity can reveal their underlying capacity for consciousness, highlighting EEG as a powerful prognostic tool[11].

Many EEG features correlate to a patient's capacity for consciousness. In an EEG signal, spectral power measures the strength of brain activity within different characteristic frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (>30 Hz). Spectral power, in particular alpha power, has emerged as a prominent feature in discerning a patient's capacity to recover consciousness[12]. Anesthesia induces a state of altered consciousness, causing a shift in power from higher frequency (beta, alpha) to lower frequency (theta, delta) spectral bands in EEG recordings[13]. However, recent research suggests that alpha power may be a marker of consciousness specific to anoxic patients, whose injuries are caused by a lack of oxygen to the brain[14]. Therefore, the need for diverse and comprehensive markers of consciousness is evident.

The Perturbational Complexity Index (PCI), currently a leading method to discern a patient's level of consciousness, measures the complexity of in-

formation in an EEG signal[15]. PCI evaluates the complexity of EEG signals after transcranial magnetic stimulation (TMS), a technique that uses magnetic pulses to stimulate the brain. This method differentiates between levels of consciousness with high accuracy[16]. However, PCI is limited by its high computational complexity and the scarce clinical availability of TMS machines, highlighting the need for accessible measures of consciousness[17,18].

We propose anesthesia as an effective perturbational mechanism to assess EEG signal complexity. This method is clinically accessible, taking advantage of existing conditions in the PICU: continuous EEG monitoring and anesthetic infusion. We aim to use these clinical conditions to explore how a patient's EEG activity changes with anesthesia—can EEG features predict recovery? Our objective is to develop a tool that uses EEG markers to assess the capacity for a return of consciousness in epileptic pediatric patients.

We analyzed various features of EEG complexity and spectral power during baseline (non-sedated) and sedated states, observing if these measures predicted the recovery of consciousness. Baseline states were taken when the patient was not experiencing active seizures, either before or after sedation depending on the availability of recordings. We further considered the prognostic potential of neural reconfiguration under anesthesia, looking at the difference in EEG activity between the baseline and sedated states. Our goal is to extend findings in adults to pediatric patients, exploring whether EEG markers of consciousness can predict functional outcomes. We hypothesize that worse outcomes will be correlated with lower neural complexity, aligning with the findings that high complexity supports consciousness[3–5,10]. We aim to find accessible prognostic markers in epileptic pediatric patients to improve clinical outcomes and advance the understanding of consciousness recovery in the PICU.

Methods

Participant Selection

Participants were selected from an EEG dataset of 41 pediatric patients with a variety of brain injuries, collected in the McMaster Children's Hospital PICU. These patients were given sedatives at different dosages, including midazolam, propofol, and fentanyl. To observe meaningful neural trends, we only used sedation recordings known to alter EEG activity: propofol and midazolam. Propofol and midazolam act through GABA-A receptors, increasing the effects of the inhibitory neurotransmitter GABA to reduce brain activity[19,20]. Fentanyl is used for pain management and does not suppress brain activity to the same extent as propofol and midazolam. Therefore, a recording containing only fentanyl was considered a baseline state.

Patients' three-month outcomes were measured with the Glasgow Outcome Scale-Extended (GOS-E), a scale from 1 (patient death) to 8 (upper good recovery) used to assess functional outcomes[21]. Our inclusion criteria were [1] epilepsy etiology, [2] over five years old, [3] available baseline and sedation recording, [4] available three-month GOS-E measurement, and [5] good quality EEG data. After selection, twelve participants remained.

The demographics of the participants are as follows: six females and six males with a mean age of 11.0±2.2 years (range: 8.0-14.0 years), ten participants with midazolam infusion and two with propofol infusion, seven participants with good recovery (GOS-E $\geq$ 7) and five with poor recovery (GOS-E < 7) recorded three months post-injury. We defined a GOS-E threshold of $\geq$ 7 as the recovered group and < 7 as non-recovered.

Upon admission to the PICU, eight participants presented with status epilepticus, an emergency characterized by long or repeated seizures. Of the remaining four participants, three presented with acute seizures, but it is unclear whether they met the criteria for status epilepticus. The final participant had systemic injuries related to seizure events. Five participants had a previously established diagnosis of epilepsy, whereas seven participants were either new onset or had an uncertain epilepsy history. The majority of patients likely have generalized seizures, affecting large parts of the brain. However, two participants presented with features that may suggest focal epilepsy: one participant with Sturge-Weber syndrome, commonly associated with focal seizures, and another participant with a temporal-parietal abscess who also likely had focal seizures.

Sedation Considerations

Each participant's sedation recording is within 50 hours of their baseline recording. Propofol and midazolam have different effects on EEG, particularly within the beta frequency band that is associated with fast, short wavelength EEG frequencies[22,23]. Midazolam can increase beta spectral power, contrasting with reduced beta spectral power observed with propofol sedation. However, at lower drug doses (sedative doses), propofol and midazolam have similar effects on EEG activity. As mentioned, both drugs act on the same receptor to yield neural inhibition. Therefore, we included participants with either midazolam or propofol sedation to observe if general neuronal inhibition, irrespective of the specific drug mechanism, could predict the recovery of consciousness.

A high sedative dose was preferentially chosen for study inclusion in participants with multiple sedation recordings, as it yields more significant neural change. Infusions were within a typical PICU sedation maintenance range[13]. Midazolam doses ranged from 1-6 µg/kg/min, and propofol doses were between 1-4 mg/kg/hr. Most participants had a dose on the higher end of these ranges, and the doses were similar between the recovered and non-recovered groups. While the participant with the propofol infusion of 4 mg/kg/hr fell slightly outside of the standard sedative range, this variation did not meaningfully impact the results. The specific dose selected for each participant is noted in Supplementary Table 1.

EEG Preprocessing

EEG data was recorded with a 26-channel system, where 26 electrodes were placed around participants' scalps to record whole-brain electrical activity. These electrodes were placed in the standard 10-20 system arrangement that distributes electrodes across the scalp. This data was collected in the McMaster Children's Hospital PICU. Recordings ranged between 7-30 minutes long. Data was preprocessed prior to further analysis, a common EEG technique to remove non-physiological noise. Preprocessing generally contains the same standard steps: filtering out extremely high and low-frequency data (known to be non-physiological), selecting a reference electrode whose activity is used as a baseline for other electrodes, and fragmenting the EEG recording into shorter segments (epochs) for a more fine-tuned analysis. An example of our EEG data segmented into epochs is shown in Figure 1B, where each horizontal line indicates the data from one electrode, and the dotted vertical lines isolate 10-second data stretches (our epoch length).

Our preprocessing pipeline used Python tools (MNE, NumPy, Pandas, SciPy, AutoReject[24]). We performed Independent Component Analysis (ICA), a technique to find and remove EEG components related to eye and heart activity that could introduce electrical noise into the recordings. Examples of components found with ICA are shown in Figure 1A, where red indicates positive activation and blue indicates negative activation. Scalp regions that have more intense colour have higher activity. A component caused by an eye blink will have high activity in the frontal electrodes (near the "nose" in the scalp schematic in Figure 1A). Components originating from non-brain sources, such as eye blinks, were identified and removed

before reconstructing the EEG signal. We then used the EEG-specialized machine learning tool AutoReject to automatically detect, interpolate, and remove particularly noisy epochs.

We filtered out low and high frequencies to yield recordings with activity between 0.5-45 Hz. We further applied a notch filter, a filter used to drastically decrease the strength of a signal at a target frequency, at 60 Hz. 60 Hz corresponds to the AC electrical power of North America, appearing as significant electrical noise in the EEG signal. The Mastoid electrodes are located near the ear and do not generally record high levels of brain activity. Therefore, we set the Mastoid electrodes as references, with their activity serving as a baseline from which to analyze other electrodes.

We inspected each recording's Power Spectral Density (PSD) plot, which shows how power (y-axis) is scattered across frequencies (x-axis). Example PSD plots are shown in Figure 1C and Figure 1D. A high-power value at a given frequency indicates that frequency is prominent in the EEG signal. Through visual analysis of these plots, we identified many high-power peaks around 28 and 44 Hz. These peaks were likely induced by medical equipment in the PICU. The peak parameters (power, width, and frequency) were identified with the power spectrum modelling software Fitting Oscillations and One Over F (FOOOF), a Python package used to find key power and frequency characteristics of EEG signals[25]. We used a notch filter to remove the peaks at 28 and 44 Hz, adjusting the power of the filter to the size of the peak. Electrodes that contained excessive noise not removed during prior preprocessing were identified, and up to 25% of electrodes were removed per recording to improve data quality.

Results show successful noise reduction and improved signal quality post-preprocessing. The dataset contains 145 recordings from 12 participants. Four recordings failed the preprocessing pipeline due to excess noisy epochs. Within the remaining data, 4.69% of all EEG electrodes were removed.

EEG Analysis

We calculated 12 spectral power and 27 complexity features for each epoch and electrode in the EEG recordings. We then calculated the median value for each feature across all epochs and electrodes, ensuring we had only one value per feature for each participant for simplicity of analysis. Further analysis was performed with these median values to compare time-averaged EEG feature values with patient recovery. The 12 power features included the relative and absolute spectral power values for the previously mentioned characteristic frequency bands (delta, theta, alpha, beta, and gamma). Additional power features were the slope and slope offset of the PSD plots, measures of how spectral power decays with increasing frequency (see examples of this frequency-dependent decay in Figures 1C and 1D). Complexity features were a range of neural complexity and entropy features implicated in signal processing and consciousness science. Some examples are Lempel-Ziv complexity (LZC), Permutation entropy, and Shannon Entropy.

We plotted each feature's value against GOS-E with a linear regression line and calculated the Pearson correlation coefficient (r) to assess the strength of a linear relationship between two variables. We further calculated the R^2 and p-value of the Pearson correlation, which indicate whether the linear model is a good fit to the data and the significance of the observed relationship, respectively. R^2 is a value between 0 and 1, with values closer to 1 indicating a stronger fit of the model to the data. The p-value is a measure of statistical significance, where smaller p-values indicate greater significance (with $p < 0.05$ often used as a cutoff).

The R^2 value was validated using leave-one-out cross-validation. This technique evaluates the linear model's fit by sequentially removing individual data points and assessing the model's performance on the omitted data. Cross-validation allows for a more thorough analysis of the relationship between GOS-E and EEG features; a correlation with a high R^2 but low validated R^2 indicates that the correlation is well fitted to the present data but may not generalize well to new data (overfitting). We performed this linear analysis three times, looking at trends in the baseline state, sedated state, and the absolute difference between the baseline and sedated states (difference state).

Subsequently, we divided participants into recovered and non-recovered groups. Differences between recovery groups for each EEG feature were assessed with a Mann-Whitney U test, which uses the test statistic U to evaluate whether the distributions of two groups differ significantly. The Mann-Whitney U test was performed on the baseline, sedated, and difference states. A p-value of < 0.05 was used to evaluate whether the differences observed in the Mann-Whitney U test were statistically significant. Features that yielded significant U values had better separability between groups and were selected for further analysis.

These significant features were used in logistic regression, a machine-learning technique that calculates the probability of an instance belonging to one of two classification categories. We trained a logistic regression to determine whether a patient was recovered (high GOS-E) or non-recovered (low GOS-E) from EEG features. We tested different combinations of features to find those that yielded the highest predictive accuracy. After finding this combination, we modified the logistic regression to enhance performance by normalizing EEG features to the same scale and adjusting the model's hyperparameters (adjustable settings which control the model's behaviour and complexity). We found that the ideal hyperparameters were those that reduced the model's complexity, ensuring it generalizes well to new data, and used a regression algorithm well suited for small datasets. The model was validated with leave-one-out cross-validation.

The 95% confidence intervals for accuracy and the area under the Receiver Operating Characteristic (ROC) curve (AUC-ROC) were calculated to evaluate the model's performance. AUC-ROC quantifies how well a model distinguishes between two groups, with higher values indicating better discrimination. Confidence intervals provide a range of values for these metrics, indicating the model's reliability and generalizability. To estimate these intervals, we used bootstrapping—a resampling technique that repeatedly tests random subsets of the data to assess the model's reliability. We used bootstrapping with 1000 iterations to obtain a stable measure of the confidence intervals without using excessive computational resources. The significance of the accuracy and AUC-ROC was assessed with a permutation test, which evaluates whether the results are statistically meaningful by randomly shuffling data and recalculating outcomes over 1000 iterations.

Many participants have several EEG recordings with different sedative doses. As previously discussed, we selected the highest available sedative dose within the typical PICU sedation range for study inclusion. To validate our dosage selection method and the logistic regression's performance, we ran a series of logistic regressions with randomized sedative EEG recordings selected from a participant's available recordings. This analysis helps reveal whether higher sedative doses are optimal for EEG investigation. For each participant, a random recording within the typical PICU sedative range was selected. The range of available sedation doses for each participant is noted in Supplementary Table 1. These recordings were then subject to logistic regression with the same parameters as previously outlined, and the accuracy and AUC-ROC were assessed. We repeated this randomization 1000 times and calculated the average accuracy and AUC-ROC across all trials. This additional analysis helps to better understand the prognostic capabilities of anesthetic infusion, regardless of the selected dosage.

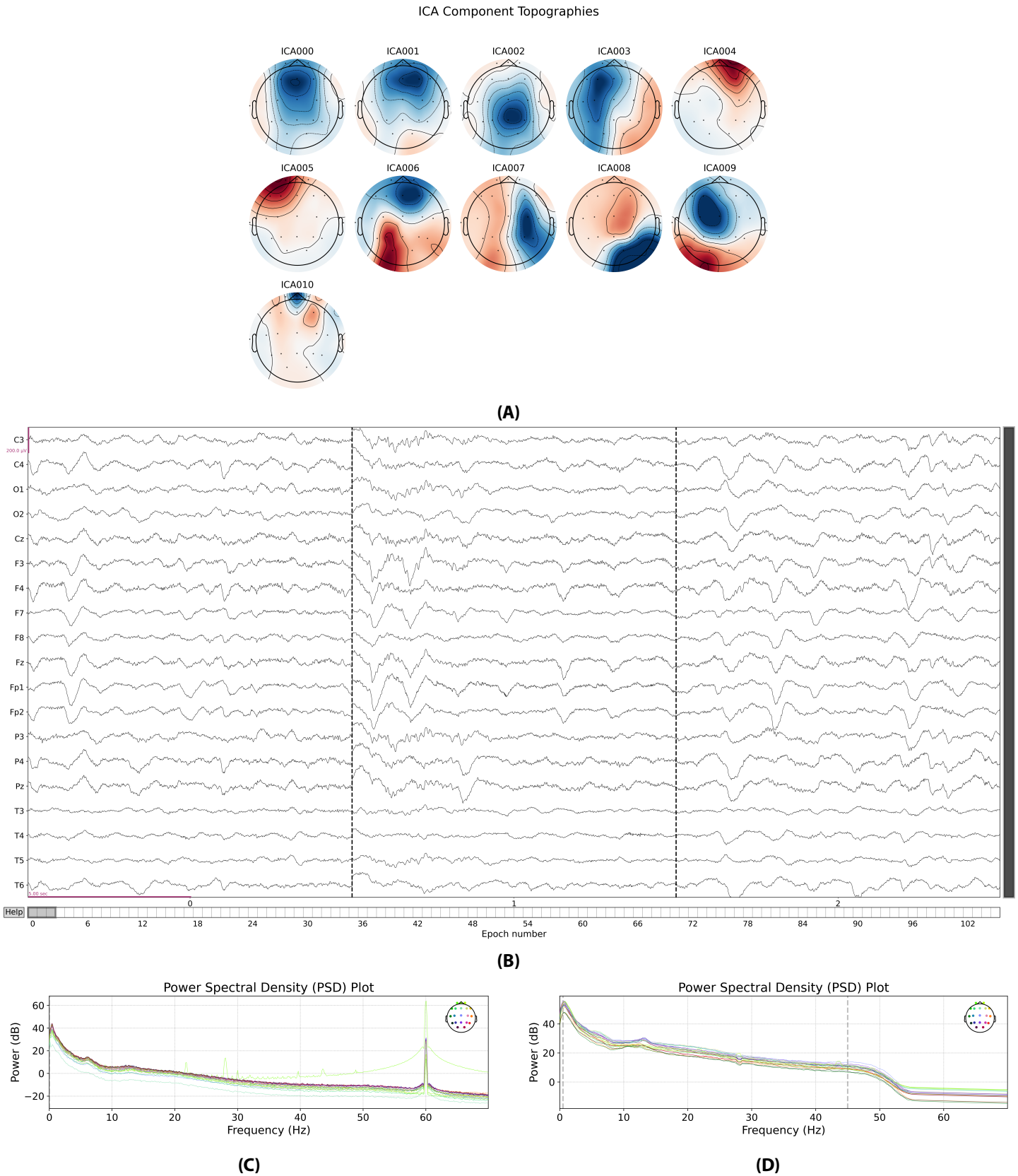


Figure 1. Visualization of the preprocessing pipeline for an example EEG recording. (**A**) Independent Component Analysis (ICA) to identify different components of activity across the scalp. Red regions represent positive activation, and blue regions represent negative activation. Higher opacity indicates greater activity in that area. (**B**) Time series of EEG activity segmented into epochs, showing multi-channel brain activity. The horizontal axis represents different 10-second epochs, while the vertical axis shows different EEG electrodes. Each line shows how voltage in a given electrode varies over time. (**C**, **D**) Power Spectral Density (PSD) plots of the example EEG recording before (left) and after (right) preprocessing. The x-axis represents frequency (Hz), and the y-axis shows spectral power (dB). Notably, preprocessing reduced the large peak of noise at 60 Hz.

Results

Baseline Analysis

None of the power or complexity features calculated from the baseline (non-sedated) recordings were significant linear predictors of participants' GOS-E scores. One complexity feature, Permutation LZC, yielded significant differences between low and high GOS-E groups (U=30.0, p=0.048). Permutation LZC is a complexity feature that looks at nonlinear signal variability in an EEG signal[26]. We did not find any other results of interest in the baseline state, so we did not pursue further analysis of these recordings.

Difference Analysis

The perturbation of the brain under anesthesia (absolute difference between the baseline and sedated state) yielded several significant linear trends. Although none of the power features were significant linear predictors of participants' GOS-E scores, we found seven complexity features that were significantly correlated with GOS-E, including Approximate entropy (a measure of pattern predictability designed for physiological data) and Shannon entropy (a foundational entropy metric that assesses the amount of information in a signal)[27]. All linear trends had a negative coefficient, indicating that smaller changes in complexity between baseline and sedation states are predictive of higher GOS-E. While many of these features yielded strong linear trends, their cross-validated R^2 values were poor, indicating that the observed trends do not generalize well to new data.

There were significant differences between low and high GOS-E participants' alpha power using a Mann-Whitney U test (U=30.0, p=0.048). We found two complexity features that yielded significant GOS-E group differences. These results demonstrate that the perturbation of the brain under anesthesia may predict patient recovery. While the complexity features had a negative relationship with GOS-E (i.e., a smaller change in complexity correlated with a higher GOS-E), alpha power had a positive trend; greater differences in alpha power between baseline and sedation were associated with greater recovery.

Sedation Analysis

The most significant results, both linearly and with group differences, are in the sedated state. Twelve complexity features had significant Pearson correlations with participants' GOS-E scores. Most of these trends persisted after cross-validation, indicating good generalizability. Twelve complexity features and three power features yielded significant differences between low and high GOS-E groups. These included relative and absolute beta power, relative delta power, Permutation entropy, and LZC among others (see Supplementary Tables 2 and 3 for more examples).

Specifically, complexity features rate entropy and Lopez-Ruiz-Mancini-Calbet Complexity (HC-LMC) separated all participants on a single-subject level between recovered and non-recovered groups. Rate entropy and HC-LMC had strong linear trends and high cross-validated R^2 values of 0.503 and 0.566, respectively, indicating they are highly predictive of patient recovery. Rate entropy positively correlated with good recovery, whereas HC-LMC negatively correlated with recovery. These two relationships are shown in Figure 2, and violin plots of the features separated into GOS-E groups are shown in Figure 3. Violin plots provide a visual representation of the distribution and variability of rate entropy and HC-LMC across GOS-E groups, highlighting differences between recovered and non-recovered patients. For example, in Figure 3, we see differences in rate entropy and HC-LMC for low and high GOS-E groups, which express an inverse relationship between these features.

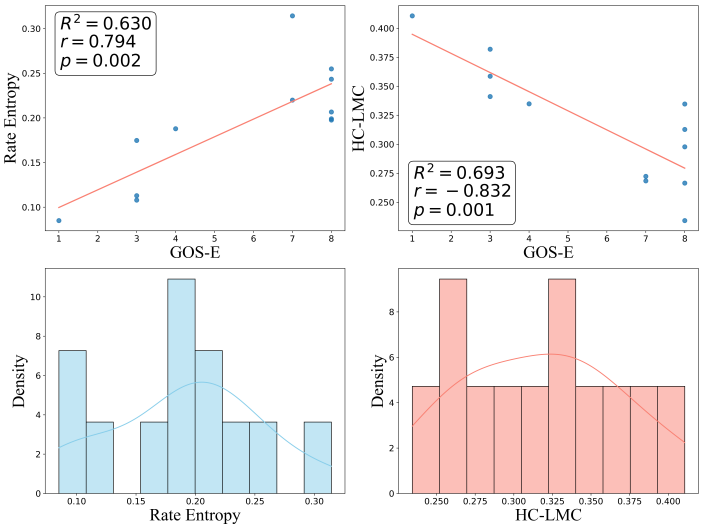


Figure 2. Relationships between EEG complexity features and patient recovery under sedation. (Top Left) Rate entropy vs. GOS-E. A strong positive correlation indicates that higher entropy rates are associated with better recovery. (Top Right) Lopez-Ruiz-Mancini-Calbet Complexity (HC-LMC) vs. GOS-E. A strong negative correlation suggests that lower HC-LMC values are predictive of better recovery. Pearson's correlation coefficient, line goodness-of-fit, and the linear relationship's significance are shown for each scatter plot. Histograms showing the distribution of each metric across patients are shown under the scatter plots.

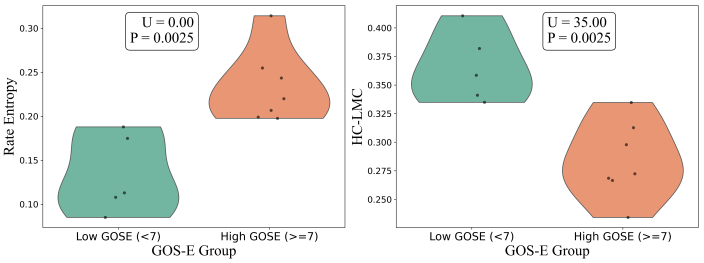


Figure 3. Comparison of EEG complexity features between patients with poor (GOS-E < 7) and good (GOS-E ≥ 7) recovery outcomes. (Left) Violin plot of rate entropy in the two recovery groups. Patients with better outcomes exhibit significantly higher Rate Entropy values (Mann-Whitney U=0.00, p=0.0025). (Right) Violin plot of HC-LMC across recovery groups. Patients with better outcomes show lower HC-LMC values (Mann-Whitney U=35.00, p=0.0025).

Logistic Regression

We found that the combination of HC-LMC and rate entropy, recorded with sedation, gave the highest accuracy in logistic regression. The model correctly classified 100% of cases, with an AUC-ROC of 1.00 (95% CI: 1.00-1.00). The model performed worse with additional EEG features, indicating that the model favours simplicity over complexity. The permutation test p-value was 0.0010 for AUC-ROC and 0.0020 for accuracy, both highly significant results. These results suggest that the observed separability of low and high GOS-E with HC-LMC and rate entropy is unlikely to be due to chance.

Using the same logistic regression hyperparameters with HC-LMC and rate entropy as features, we randomized the sedation dose for each participant. After repeating the logistic regression 1000 times with random sedative doses, the model correctly classified an average of 80.33% of cases (95% CI: 70.72%-80.95%), with an AUC-ROC of 0.777 (95% CI: 0.769-0.785).

Discussion

We explored the prognostic potential of various EEG features of spectral power and complexity for both baseline (non-sedated) and sedated states, considering whether baseline, sedation, or the difference between the two states' features (due to neural reconfiguration) was most predictive of recovery. We found that the baseline state has little prognostic power, and the difference state had some predictive capability. However, the linear trends seen in the difference state did not persist after validation, indicating they may result from overfitting. We found some ability to discriminate between low and high GOS-E in the difference state, suggesting that changes in the brain with anesthesia may predict the recovery of consciousness. However, more work is needed to validate these trends in larger sample sizes.

Overwhelmingly, we found the most substantial prognostic power in the sedated state. Many EEG complexity features and a few power features were highly predictive of a participant's recovery. These trends persisted after cross-validation, indicating they are reliable and may generalize to other epileptic patients. Rate entropy and HC-LMC had high predictive abilities, separating participants into recovered and non-recovered groups with 100% accuracy. These results suggest that rate entropy and HC-LMC are exceptionally indicative of an epileptic patient's capacity for consciousness and may serve as powerful prognostic tools in a PICU setting.

Even with dose randomization, we found good discriminative ability between low and high GOS-E. This result indicates that measuring EEG complexity under sedation, regardless of the sedative dose, can help predict patient recovery. However, the accuracy of the logistic regression was greater with a high sedative dose. Therefore, while any sedative infusion within the typical PICU dosage range has prognostic capabilities with EEG complexity analysis, higher sedative doses best predict recovery.

The complexity feature rate entropy is the time-resolved derivative of entropy, a measure of the amount of information, or uncertainty, contained in a signal[28]. We demonstrated that higher rates of change correlate to high GOS-E, indicating that greater changes in entropy are associated with better recovery. The brain reconfigures with injury; post-injury plasticity is a compensatory mechanism that aids recovery[29,30]. Epilepsy and neuroplasticity are fundamentally interconnected—epilepsy can induce plasticity, and the extent of neural reconfiguration influences epileptic progression[31,32]. Rate entropy may reflect the brain's ability to reorganize and adapt, serving as a proxy for epilepsy-induced neural plasticity. A brain with greater entropic changes after injury may be more resilient, modifying its activity patterns as an adaptive mechanism. High versatility and capacity for change may therefore underlie a patient's recovery of consciousness. These results suggest that measuring EEG changes in entropy in patients undergoing anesthesia can accurately predict their epileptic progression.

HC-LMC is a complexity feature that assesses the amount of structured patterns within a signal[33]. At low values of HC-LMC, a system is either highly ordered or disordered, whereas higher values yield a greater balance of order and disorder[34]. Interestingly, better recovery is correlated with lower values of HC-LMC—fewer structured patterns. Under anesthesia, neural complexity is lessened. Patients who maintained a balance of order and disorder under anesthesia had a lower capacity for consciousness, so the ability to sustain asymmetry of these measures with sedation may be crucial for recovery. Epilepsy induces bursts of brain activity and structured patterns in the signal, which are associated with higher HC-LMC. Therefore, participants with greater HC-LMC may have more epileptic activity and worse recovery. Although we demonstrated that HC-LMC has great potential to predict the recovery of consciousness, it is an understudied complexity feature in clinical contexts. These results highlight the need for future research to understand how HC-LMC reflects brain dynamics and to assess its utility as a predictive marker of clinical outcomes.

We demonstrated that rate entropy and HC-LMC can predict the recovery of epileptic pediatric patients in this cohort with 100% accuracy. These are not simple features of neural complexity, but intricate measures of the rate of entropic change and the balance of order and disorder. Accordingly, higher complexity is not always associated with better recovery—we observed a negative trend between GOS-E and HC-LMC. Our results demonstrate that EEG features under sedation can reveal a patient's capacity for consciousness, building on current findings. The high accuracy of the logistic regression highlights EEG as a powerful and accessible tool to predict epilepsy outcomes. EEG markers of consciousness, such as rate entropy and HC-LMC, could enhance clinical decision-making and improve our understanding of the neural mechanisms underlying recovery.

Limitations and Future Directions

While the results suggest that rate entropy and HC-LMC have strong prognostic ability, these conclusions are limited by the small sample size. We tested 12 participants; more work is needed to verify if the observed trends generalize to larger samples. Further, we had unbalanced groups: 5 participants with low GOS-E and 7 participants with high GOS-E. We had no participants with a GOS-E score of 2, 5, or 6. To substantiate these findings, our results should be validated with a greater diversity of GOS-E scores, particularly in the worse recovery group.

We showed that rate entropy and HC-LMC are powerful markers in the prognosis of epileptic pediatric patients. Future studies could investigate these markers in other etiologies—are they specific to epilepsy, or do they predict recovery in different conditions? Further, it would be valuable to explore these EEG features in a healthy control pediatric sample. These results could serve as a baseline for evaluating a patient's degree of altered neural activity.

Conclusion

We established rate entropy and HC-LMC as accurate and valuable EEG complexity features for predicting recovery in epileptic pediatric patients in the PICU. These features give insight into dynamic changes in entropy and the brain's intricate balance between order and disorder. Our findings reveal that the sedated state is the most predictive of recovery, highlighting the prognostic value of EEG analysis under anesthesia. Rate entropy and HC-LMC are clinically relevant features that perfectly classify recovery outcomes. They may reflect fundamental neural mechanisms, such as plasticity and adaptability, that are crucial for a patient's capacity for consciousness. The integration of these features into clinical decision-making could advance accessible prognostication, offering clinicians effective and accurate tools to assess recovery in epileptic patients.

Acknowledgements

This project was supported by the NSERC USRA grant. Thank you to Mark Greenberg and Kevin Jones, who collected the McMaster Children's Hospital EEG recordings. This project would not have been possible without their fundamental contributions.

Supplementary Material

Supplementary material referenced in the text of this article may be found online at https://doi.org/10.26443/msurj.v1i1.288.

References

1. Mertiri, L., Rossi, A., Huisman, L. M. & Huisman, T. A. in *Diseases of the Brain, Head and Neck, Spine 2024-2027: Diagnostic Imaging* (eds Kubik-Huch, R. A. & Roos, J. E.) 143–156 (Springer, 2024). https://doi.org/10.1007/978-3-031-50675-8_14
2. Albertini, F., Bresson, V., Tardieu, S., Milh, M. & Chabrol, B. Pediatric emergency room visits for neurological conditions: Description and use of pediatric neurologist advice. *Arch. Pediatr.* **27,** 416–422 (2020). https://doi.org/10.1016/j.arcped.2020.09.005
3. Tononi, G. & Edelman, G. M. Consciousness and complexity. *Science* **282,** 1846–1851 (1998). https://doi.org/10.1126/science.282.5395.1846
4. Schartner, M. et al. Complexity of multi-dimensional spontaneous EEG decreases during propofol induced general anaesthesia. *PLOS ONE* **10,** e0133532 (2015). https://doi.org/10.1371/journal.pone.0133532
5. Varley, T. F. et al. Consciousness & brain functional complexity in propofol anaesthesia. *Sci. Rep.* **10,** 1018 (2020). https://doi.org/10.1038/s41598-020-57695-3
6. Rasulo, F. A. et al. Processed electroencephalogram-based monitoring to guide sedation in critically ill adult patients: recommendations from an international expert panel-based consensus. *Neurocrit. care* **38,** 296–311 (2023). https://doi.org/10.1007/s12028-022-01565-5
7. Egbuta, C. & Mason, K. P. Current state of analgesia and sedation in the pediatric intensive care unit. *J. Clin. Med.* **10,** 1847 (2021). https://doi.org/10.3390/jcm10091847
8. Maschke, C., Duclos, C. & Blain-Moraes, S. Paradoxical markers of conscious levels: effects of propofol on patients in disorders of consciousness. *Front. Hum. Neurosci.* **16,** 992649 (2022). https://doi.org/10.3389/fnhum.2022.992649
9. Gunawardena, S., Chikkannaiah, M., Stolfi, A. & Kumar, G. Utility of electroencephalogram in the pediatric emergency department. *Am. J. Emerg. Med.* **54,** 26–29 (2022). https://doi.org/10.1016/j.ajem.2022.01.045
10. El Youssef, N. et al. Consciousness alteration in focal epilepsy is related to loss of signal complexity and information processing. *Sci. Rep.* **12,** 22276 (2022). https://doi.org/10.1038/s41598-022-25861-4
11. Azabou, E. et al. Value and mechanisms of EEG reactivity in the prognosis of patients with impaired consciousness: a systematic review. *Crit. Care* **22,** 1–15 (2018). https://doi.org/10.1186/s13054-018-2104-z
12. Pauli, R., O'Donnell, A. & Cruse, D. Resting-state electroencephalography for prognosis in disorders of consciousness following traumatic brain injury. *Front. Neurol.* **11,** 586945 (2020). https://doi.org/10.3389/fneur.2020.586945
13. Sarasso, S. et al. Consciousness and complexity during unresponsiveness induced by propofol, xenon, and ketamine. *Curr. Biol.* **25,** 3099–3105 (2015). https://doi.org/10.1016/j.cub.2015.10.014
14. Colombo, M. A. et al. Beyond alpha power: EEG spatial and spectral gradients robustly stratify disorders of consciousness. *Cereb. Cortex* **33,** 7193–7210 (2023). https://doi.org/10.1093/cercor/bhad031
15. Casali, A. G. et al. A theoretically based index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med.* **5,** 198ra105–198ra105 (2013). https://doi.org/10.1126/scitranslmed.3006294
16. Sinitsyn, D. O. et al. Detecting the potential for consciousness in unresponsive patients using the perturbational complexity index. *Brain Sci.* **10,** 917 (2020). https://doi.org/10.3390/brainsci10120917
17. Comolatti, R. et al. A fast and general method to empirically estimate the complexity of brain responses to transcranial and intracranial stimulations. *Brain stimul.* **12,** 1280–1289 (2019). https://doi.org/10.1016/j.brs.2019.05.013
18. Maschke, C. et al. Critical dynamics in spontaneous EEG predict anesthetic-induced loss of consciousness and perturbational complexity. *Commun. Biol.* **7,** 946 (2024). https://doi.org/10.1038/s42003-024-06613-8
19. Shin, D. J. et al. Propofol is an allosteric agonist with multiple binding sites on concatemeric ternary GABAA receptors. *Mol. Pharmacol.* **93,** 178–189 (2018). https://doi.org/10.1124/mol.117.110403
20. Eom, W. et al. The effects of midazolam and sevoflurane on the GABA A receptors with alternatively spliced variants of the $\gamma 2$ subunit. *Korean J. Anesthesiol.* **60,** 109–118 (2011). https://doi.org/10.4097/kjae.2011.60.2.109
21. Wilson, L. et al. A manual for the glasgow outcome scale-extended interview. *J. Neurotrauma* **38,** 2435–2446 (2021). https://doi.org/10.1089/neu.2020.7527
22. Miyake, W. et al. Electroencephalographic response following midazolam-induced general anesthesia: relationship to plasma and effect-site midazolam concentrations. *J. Anesth.* **24,** 386–393 (2010). https://doi.org/10.1007/s00540-010-0907-4
23. Numan, T. et al. Resting state EEG characteristics during sedation with midazolam or propofol in older subjects. *Clin. EEG Neurosci.* **50,** 436–443 (2019). https://doi.org/10.1177/1550059419838938
24. Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F. & Gramfort, A. Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage* **159,** 417–429 (2017). https://doi.org/10.1016/j.neuroimage.2017.06.030
25. Donoghue, T. et al. Parameterizing neural power spectra into periodic and aperiodic components. *Nat. Neurosci.* **23,** 1655–1665 (2020). https://doi.org/10.1038/s41593-020-00744-x
26. Bai, Y., Liang, Z., Li, X., Voss, L. J. & Sleigh, J. W. Permutation Lempel–Ziv complexity measure of electroencephalogram in GABAergic anaesthetics. *Physiol. Meas.* **36,** 2483 (2015). https://doi.org/10.1088/0967-3334/36/12/2483
27. Delgado-Bonal, A. & Marshak, A. Approximate entropy and sample entropy: A comprehensive tutorial. *Entropy* **21,** 541 (2019). https://doi.org/10.3390/e21060541
28. Shannon, C. E. A mathematical theory of communication. *Bell Labs Tech. J.* **27,** 379–423 (1948). https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
29. Nudo, R. J. Recovery after brain injury: mechanisms and principles. *Front. Hum. Neurosci.* **7,** 887 (2013). https://doi.org/10.3389/fnhum.2013.00887
30. Su, Y. S., Veeravagu, A. & Grant, G. in *Translational Research in Traumatic Brain Injury* (eds Laskowitz, D. & Grant, G.) 163 (CRC Press, 2016).
31. Jarero-Basulto, J. J. et al. Interactions between epilepsy and plasticity. *Pharmaceuticals* **11,** 17 (2018). https://doi.org/10.3390/ph11010017
32. Sutula, T. P. Mechanisms of epilepsy progression: current theories and perspectives from neuroplasticity in adulthood and development. *Epilepsy Res.* **60,** 161–171 (2004). https://doi.org/10.1016/j.eplepsyres.2004.07.001
33. Lopez-Ruiz, R., Mancini, H. L. & Calbet, X. A statistical measure of complexity. *Phys. Lett.* **209,** 321–326 (1995). https://doi.org/10.1016/0375-9601(95)00867-5

34. López-Ruiz, R. Shannon information, LMC complexity and Rényi entropies: a straightforward approach. *Biophys. Chem.* **115,** 215–218 (2005). https://doi.org/10.1016/j.bpc.2004.12.035

Submitted: 06/24/2024
Accepted: 03/06/2025
Published: 03/31/2025

Research Article

[1]Genomic Instability in Development and Disease, European Institute for the Biology of Aging, University of Groningen, Groningen, Netherlands

Keywords

Lymphoma, Chromosomal instability, IL-6 signaling, Inflammatory response, cGas/STING

Email Correspondence

v.muller.5@student.rug.nl

https://doi.org/10.26443/msurj.v20i1.246

Veronika Müller[1]

Chromosomal Instability and IL6-Mediated Inflammatory Signaling in Human Lymphoma Cells

Abstract

Chromosomal Instability (CIN) is a hallmark of cancer that significantly impacts tumor evolution and prognosis in lymphomas. This study investigates the relationship between CIN and inflammatory signaling, particularly the IL6 pathway, in human lymphoma cell lines. Genomic analysis revealed diverse mutational landscapes across cell lines, reflecting varying degrees of genomic instability. qPCR analysis demonstrated consistent upregulation of IL6 expression upon Vincristine treatment across all tested cell lines. IncuCyte analysis of cell viability after treatment with the CIN-inducing agent BAY1217389 revealed dose-dependent effects, with certain levels of CIN potentially enhancing cell growth. These findings suggest that CIN activates IL6-mediated inflammatory signaling in lymphoma cells, contributing to a pro-tumor microenvironment and potentially modulating therapeutic outcomes.

Introduction

Lymphomas are a heterogeneous group of malignancies characterized by the clonal expansion of lymphocytes and marked genomic disruption. Chromosomal Instability (CIN), a hallmark of cancer, is particularly prevalent in lymphomas due to the physiological DNA breaks and translocations that occur in lymphocytes. These physiological processes encompass V(D)J (variable–diversity–joining) recombination[1], class switch recombination, and somatic hypermutation. While these rearrangements are tightly regulated under normal conditions, they can result in errors, leading CIN[1]. The stability of chromosomes is ensured by the physiological activity of genome maintenance pathways, like cell cycle checkpoints, DNA repair mechanisms, and the coordinated activity of the cell division machinery. These areas of DNA processing, including recombination, repair, and replication can be vulnerable to deficiencies in chromosomal stability. As a result, defects in these pathways can cause CIN[2].

CIN substantially impacts tumor evolution by dysregulating the immune response and promoting cell-intrinsic inflammatory signaling. The origins of cells with CIN-phenotype and micronuclei stem from various mitotic errors[2]. During anaphase, under-replicated DNA from S phase can form ultra-fine bridges, which may break in subsequent cell cycles, resulting in micronuclei. Unrepaired double-strand breaks can lead to chromosome breakage during mitosis, potentially causing radial chromosomes with multiple centromeres. These structures can form bulky chromosomal bridges that produce micronuclei. Additionally, fully duplicated chromosomes may occasionally lag at the metaphase plate due to kinetochore-microtubule attachment issues, leading to whole chromosome micronuclei in the next cell cycle (Figure 1)[2].

The rupture of micronuclei occurs due to nuclear envelope (NE) instability[3]. Chromosomes distant from spindle microtubules recruit non-core NE proteins like nuclear pore complex, Lamin B, and Lamin B receptor, disrupting the density of these proteins and leading to NE fragility. Aurora B kinase, which senses lagging chromosomes, establishes a CDK1 phosphorylation gradient by stabilizing Cyclin B[4]. This disrupts the phosphorylation of proteins involved in nuclear envelope assembly and disassembly, further destabilizing the NE[4]. Consequently, the fragile micronuclei are prone to rupture, releasing their double-stranded DNA content into the cytoplasm and, activating the cGAS/STING-pathway (Figure 2)[5].

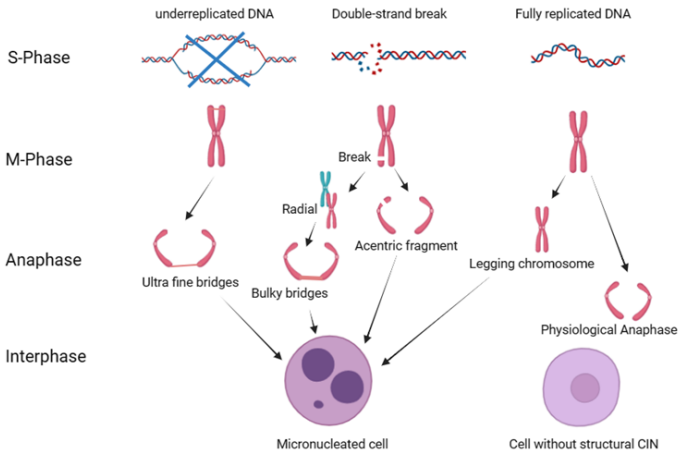


Figure 1. Schematic representation of the origins of cells with CIN-phenotype and micronuclei.

In physiological conditions, the cGAS-STING pathway detects cytoplasmic DNA. When cytoplasmic DNA binds to cGAS, it catalyzes the synthesis of cGAMP from ATP and GTP. This cGAMP acts as a second messenger, binding to STING on the ER membrane and triggering its activation. Activated STING then translocates from the ER to the Golgi apparatus, recruiting and activating TBK1 and IKK kinases. These kinases subsequently activate the IRF3 and NF-κB signaling cascades, leading to the expression of type I interferons and inflammatory cytokines that enhance immune responses[5]. However, under chronic CIN conditions, this pathway becomes

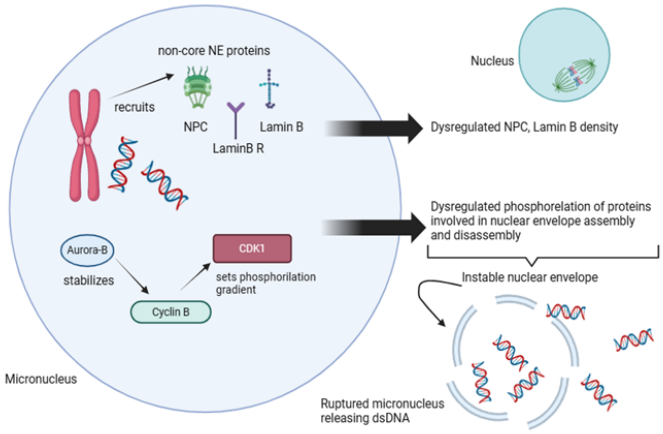


Figure 2. Schematic representation of the pathways involved in the rupture of a micronucleus as a result of an insatiable nuclear envelope (NE).

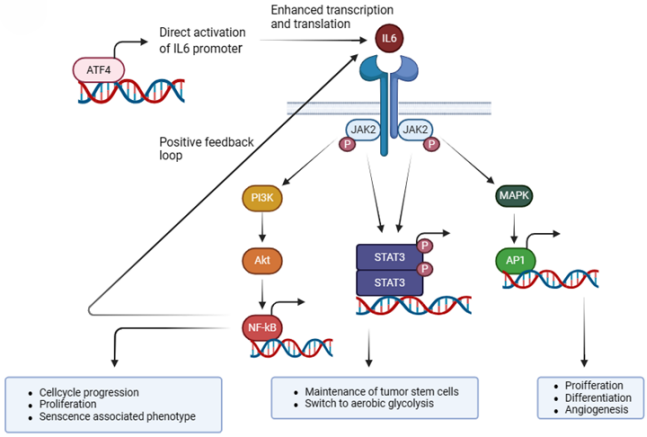


Figure 4. IL6, NF-κB and STAT3 in CIN associated tumorigenesis.

dysregulated, as illustrated in Figure 3.

The cGAS/STING pathway under chronic CIN leads to immunosuppression, a pro-tumor microenvironment, and metastasis[5]. Cytoplasmic DNA, released from ruptured micronuclei, is detected by cGAS, which catalyzes cGAMP formation and activates STING[5]. Chronic activation of this pathway results in persistent genomic instability and cellular stress, potentially overwhelming cellular machinery and leading to ER stress and unfolded protein response (UPR) activation[6]. Cancer cells exploit the UPR to adapt to the hostile tumor microenvironment, promoting cell survival and therapy resistance[7]. Prolonged activation of STING signaling may trigger negative feedback mechanisms, including epigenetic modifications that downregulate STING expression[8], ultimately resulting in immunosuppression and a pro-tumor microenvironment (Figure 3).

IL6, NF-κB, and STAT3 play crucial roles in CIN-associated tumorigenesis. ATF4, downstream of PERK, directly activates the IL6 promoter[9]. IL6 signaling promotes tumor growth, metastasis, and therapy resistance by activating multiple pathways. The PI3K/AKT pathway, activated by IL6, promotes cell survival and growth[10]. AKT activation leads to NF-κB activation, which regulates genes associated with inflammation, cell survival, and proliferation[11]. IL6 also activates the JAK/STAT pathway, leading to STAT3 phosphorylation. Activated STAT3 maintains the stemness

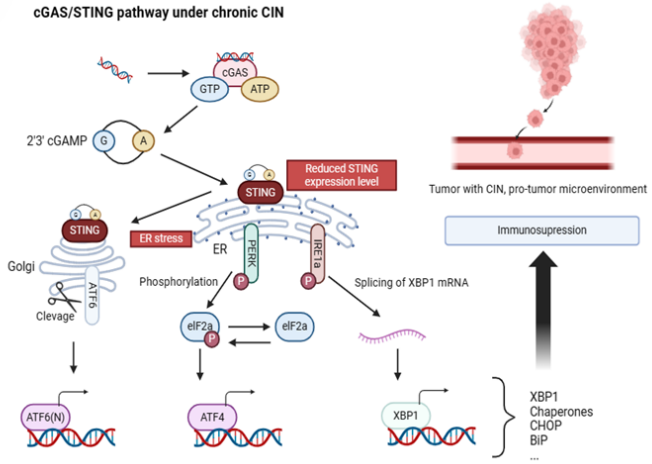


Figure 3. Schematic representation of the cGAS/STING pathway under chronic CIN, leading to immunosuppression, pro-tumor microenvironment, and metastasis.

of tumor-initiating cells and promotes metabolic reprogramming in cancer cells[12]. Additionally, IL6 activates the MAPK pathway, leading to AP-1 activation, which promotes cell proliferation, differentiation, and angiogenesis[13]. These interconnected pathways collectively contribute to tumor growth and metastasis in the context of CIN (Figure 4).

A recent study by Hong et al. highlighted the role of CIN in driving inflammatory signaling pathways, particularly the IL6/STAT3 axis, in breast cancer and a Lymphoma mouse model[5]. However, the specific implications of this pathway in human lymphoma cells remain unclear. This gap in knowledge represents a critical area for investigation, given the established link between CIN and prognosis/treatment response in lymphomas.

This study aims to investigate whether CIN in human lymphoma cells activates the IL6-mediated inflammatory signaling pathway and influences cell viability, potentially contributing to a pro-tumor microenvironment that modulates therapeutic outcomes. To address this question, a multi-faceted approach was employed combining genomics analysis, quantitative assessment of inflammatory cytokine expression, and evaluation of cell viability under CIN-inducing conditions.

This research utilizes multiple lymphoma cell lines, inducing CIN through inhibition of Monopolar spindle 1 (MPS1) kinase. The proliferative response to CIN induction and levels of key cytokines, including IL6 and its receptor was analyzed, using qPCR. Additionally, the response of these cells to Vincristine, a common chemotherapeutic agent, was assessed to infer possible implications for therapy modulation[5].

Understanding the relationship between CIN and inflammatory signaling in lymphoma is crucial for developing more effective treatment strategies. By elucidating these mechanisms, we may identify novel therapeutic targets and improve patient outcomes.

Methods

Genomics Analysis of Publicly Available Datasets

To investigate the molecular and genetic landscape in human lymphoma cell lines and set a foundation for cell line comparison, a genomics analysis of publicly available datasets was performed. Genomic data for the lymphoma cell lines Daudi, Ramos, SU-DHL-10, SUDHL-6, OCI-LY3 was downloaded from the cBioportal for Cancer genomics. To analyze datasets, we examined the data for mutation count and fraction of the genome altered, as well as specific oncogenes and genes significantly altered in the

lymphoma cell lines.

Cell Culture

The cell lines used were selected based on their relevance to the biological context under investigation. The inclusion of multiple cell lines enables a comparative analysis, ensuring that observed effects are not cell-type specific but rather reflect broader principles. These models were chosen based on their established use in previous studies.

The human lymphoma cell lines Ramos (Human Burkitt lymphoma cell line, hypodiploid karyotype, 45 Chromosomes), Daudi (Human Burkitt lymphoma cell line, nearly diploid karyotype with 66% stability of the cells within their stem line, 20% polyploid), SU-DHL-6 (B-cell non-Hodgkin lymphoma cell line, hyperdiploid, 47 Chromosomes), SU-DHL-10 (Human B-cell lymphoma, hyperdiploid karyotype), and OCI-LY3 (Human B cell lymphoma, hypertriploid karyotype) were cultured using RPMI medium with 20% fetal bovine serum, 1% penicillin streptomycin at 37 °C with 5% CO_2 for 72 hours in a humidified incubator.

qPCR

To assess the baseline and treatment induced inflammatory response, as well as the dose-dependent drug response, the human lymphoma cell lines Ramos, Daudi, SU-DHL-6, SU-DHL-10, and OCI-LY3 were tagged with mRaspberry and subsequently treated with 0nM, 2nM and 5nM of Vincristine. The dosage was established in the previous study by Hong et al., treating cancer cells with CIN[5]. A qPCR for the detection of cytokines and the IL6 receptor was performed with IL6, CCL5, CXCL10, IL-8 and IL6R primers. We used Tubulin and HPRT1 as housekeeping genes and the reaction master mix iTaq Universal SYBR Green Supermix (Bio Rad, California) for detection. Total RNAs were extracted from all cell lines using the NucleoSpin RNA Plus RNA Isolation kit (Macherey-Nagel™, UK), and total cDNAs were synthesized using the LunaScript RT Master Mix Kit (Bioke, Netherlands) in accordance with the respective manufacturer's protocol. IL6, CCL5, CXCL10, IL-8, and IL6R transcripts as well as the housekeeping genes, Tubulin and HPRT1, were detected by dye based quantitative real-time polymerase chain reaction.

Transduction for IncuCyte Analysis

To ensure that the IncuCyte analysis is performed on a homogenous population of cells expressing the desired genetic modifications at consistent levels, a transduction was performed. We transduced target Eph4 cells with lentiviral vectors encoding the genes of interest (dnMCAK, MPS1-D637A-T649A) or empty vector controls under different promoters (pGK, pUBC, pEF1a). We transfected HEK293T cells with the lentiviral transfer plasmid along with packaging plasmids psPAX2 and pMD2.G using TurboFect transfection reagent. Viral supernatants were harvested 48- and 72-hours post-transfection, filtered through a 0.45 µm filter, and supplemented with 8 µg/ml polybrene, before transducing Eph4 cells in two rounds. Two days after the second round of transduction, we selected transduced cells by replacing the medium with fresh medium containing 0.75−1.5 µg/ml puromycin until all non-transduced control cells had died. Stably transduced cell populations were then expanded under normal culture conditions without antibiotic selection. Thereafter, cells were transfected with H2B-mCherry fluorescent protein to enable cell imaging and analysis.

IncuCyte Analysis

The IncuCyte Live-Cell Analysis System was chosen for assessing cell viability and proliferation as it provides real-time, non-invasive monitoring of cell growth, allowing for the capture of temporal dynamics in response to treatment.

For the evaluation of the dose-dependent effects of BAY1217389-induced CIN on cell growth and viability in different human lymphoma cell lines, we treated Daudi, Ramos, and SU-DHL-10 with varying concentrations of the CIN-inducing Monopolar spindle 1 (MPS1) kinase inhibitor BAY1217389 (Selleckchem): 0 nM (control), 1 nM, 2.5 nM, 5 nM, 10 nM, 15 nM, 20 nM, and 25 nM. The concentrations of BAY1217389 (0-25 nM) were selected based on previously reported IC50 values in various cancer cell lines (ranging from 3.7 to 12 nM) and prior literature[1,5]. The cells were seeded in triplicate at four different initial densities: 8,000, 14,000, 20,000, and 26,000 cells per well (c/w) in a 96-well plate. The 96-well plate was placed in the IncuCyte Live-Cell Analysis System (Sartorius) and monitored for 72 hours. Phase-contrast and fluorescent images were acquired every 3 hours to track cell growth and viability. The IncuCyte software quantified the fluorescent object count, which corresponds to the number of viable cells, at each time point. We curated the obtained data and double-normalized the raw fluorescent object count data to account for potential variations in initial seeding densities and to facilitate comparison across different conditions and cell lines. The normalization process involved:

1. normalizing the raw data to the time 0 (initial seeding) count for each well;

2. normalizing the time 0-normalized data to the average of the respective vehicle control (0 nM BAY1217389) for each cell line and initial seeding density.

The normalization process allowed for comparison between vehicle control and treatment conditions as well as the observation of growth patterns. Subsequently, to gain insights into the temporal dynamics of the drug's effects and identify the time points at which the drug's anti-proliferative or cytotoxic activity is most pronounced, we performed a Pearson correlation analysis between the drug dose and the normalized cell count at each timepoint and for each seeding density. The Pearson correlation analysis provides a statistically robust method for quantifying the time-dependent relationship between BAY1217389 concentration and cell viability. This approach captures the dynamics of cellular responses to CIN induction. It reveals possible transition points where treatment effects become pronounced and identifies cell-line specific response patterns. Subsequently, a single factor ANOVA was applied to test statistical significance.

Results

Genomics Analysis of Publicly Available Datasets

Analysis of publicly available genomic datasets revealed diverse genomic landscapes across five human lymphoma cell lines: Ramos, Daudi, SU-DHL-6, SU-DHL-10, and OCI-LY3. These cell lines exhibited varying degrees of genetic alterations, providing insights into their potential biological behavior and therapeutic responses.

The selected lymphoma cell lines exhibit a range of karyotypic complexities, reflecting the genetic heterogeneity observed across lymphoma subtypes. Daudi and Ramos, for instance, share the hallmark t(8;14) IGH-MYC translocation characteristic of Burkitt lymphoma, while SU-DHL-6 and SU-DHL-10 harbor the t(14;18) IGH-BCL2 rearrangement, commonly associated with diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma. OCI-LY3, with its hypertriploid karyotype (Table 1) and concurrent IGH-SPIB and BCL amplifications, represents a more genetically complex model.

OCI-LY3, derived from anaplastic large-cell lymphoma, demonstrated the

Table 1. Genetic Landscape of the investigated human lymphoma cell lines (Leibniz Institute DSMZ: Details, https://www.dsmz.de/)

Cell line	Cytogenetics
Daudi	human near diploid karyotype with 20% polyploidy - 46(45-48)<2n>XY/XXY, +7, -9, t(8;14)(q24;q32)
Ramos	45(44-46)<2n>X, -Y, del(2)(p16p23), t(8;14)(q24;q32.2), add(14)(p11), der(16)t(7;16)(q11.2;p13.2), der(17)t(13;17)(q33;p11.3)t(13;3)(q13;q?25), del(20)(q12); carries t(8;14) effecting IGH-MYC rearrangement
SU-DHL-6	human hyperdiploid karyotype with 2% polyploidy - 47(42-48)<2n>X, -Y, +6, +7, del(4)(q23), del(6)(p21.3p22.2), i(6p), del(7)(q?22q?32), der(8)t(8;9)(q24;p13), der(9)t(8;19;9)(q24;q13;p13), dup(11)(q24q25), t(14;18)(q32;q21), der(22)t(?7;22)(?q32;p11) - sideline with dic(8;9)(q24;p13), ider(8)(q10)t(8;9(q24;p13) - matches published karyotype - carries t(14;18) effecting IGH-BCL2 fusion
SU-DHL-10	human flat-modded hyperdiploid karyotype - 47(43-48)<2n>XY, +7,der(8)t(X;8)(q25;p23)t(8;X)(q24;q26)t(X;14)(q28;q32), del(10)(q22q24), der(11)t(Y;11)(q11;q25), der(14)t(8;14)(q24;q32), der(18)t(14;18)(q32;q21) — carries concurrent rearrangements of IGH with MYC and BCL2 — resembles published karyotype
OCI-LY3	human flat-moded hypertriploid karyotype; 72-77<3n>XXYY, +1, +9, -10, +13, +14, -17, +19, +20, +22, der(1)t(1;17)(p13;q12)x2, der(4)t(4;18)(q31;q21)x2, del(6)(q13)x2, der(6)t(6;6)(p24;q12), der(7)t(6;7)(p24;p22), der(14)t(14;19)(q32;q13.3)x2, del(18)(q21), der(19)t(4;19)(q21;q13)t(4;18) (q31;q21)x2, der(19)t(14;19), dup(20)(q11q13)x2; sdl with der(6)t(6;12)(p21;q21), der(7)t(5;7)(?p15;p24) etc; resembles published karyotypes; carries cryptic t(14;19) with rearrangement of IGH and SPIB, and t(4;18) with copy number amplification of the BCL2 region

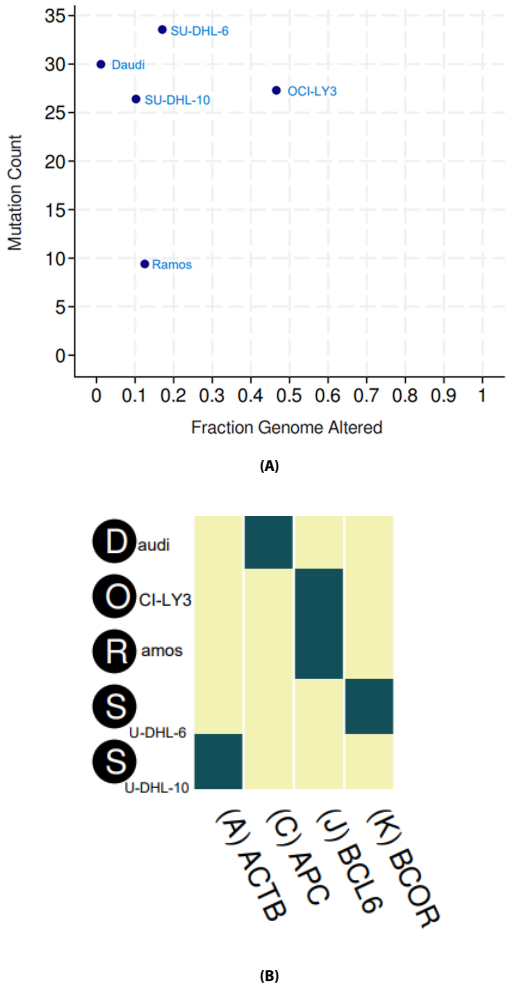


Figure 5. Genomic landscape of Ramos, Daudi, SU-DHL-6, SU-DHL-10, and OCI-LY3. (A) The mutation count is plotted versus fraction of genome altered. (B) Specific mutated genes in the respective cell lines.

most extensive history of genomic instability with a hypertriploid karyotype, complex chromosomal rearrangements, and the highest fraction of genome altered (0.461) (Figure 5A). It also carried a mutation in the BCL6 gene (Figure 5B).

SU-DHL-6, originating from B-cell non-Hodgkin lymphoma, showed a hyperdiploid karyotype with the t(14;18)(q32;q21) translocation, as depicted in Table 1, and a mutation in the BCOR gene (Figure 5A). It exhibited a high mutation count (34) and moderate fraction of genome altered (0.1679) (Figure 5B).

SU-DHL-10, another B-cell lymphoma line, displayed a hyperdiploid karyotype with rearrangements involving MYC, BCL2, and IGH loci, as well as a mutation in the ACTB gene (Figure 5A). It showed moderate levels of both mutation count (26) and fraction of genome altered (0.0970) (Figure 5B).

Daudi, characterized by a near-diploid karyotype with 20% polyploidy, carried the t(8;14)(q24;q32) (Table 1) translocation and a mutation in the APC gene. Despite a high mutation count (30), it had the lowest fraction of genome altered (0.0102) among the studied cell lines (Figure 5B).

Ramos, with a hypodiploid karyotype and the t(8;14)(q24;q32.2) translocation (Table 1), harbored an APC gene mutation (Figure 5A). It exhibited the lowest mutation count (9) but a moderate fraction of genome altered (0.1145) (Figure 5B).

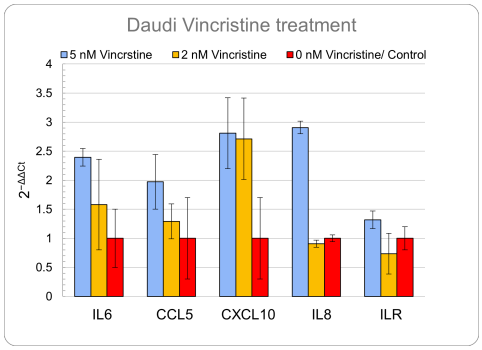
This diversity in genomic profiles across the lymphoma cell lines suggests that different molecular pathways and mechanisms may contribute to the development and progression of these malignancies. These unique genetic makeups may also influence chromosomal instability-induced inflammatory signaling pathways and responses to therapeutic interventions.

qPCR

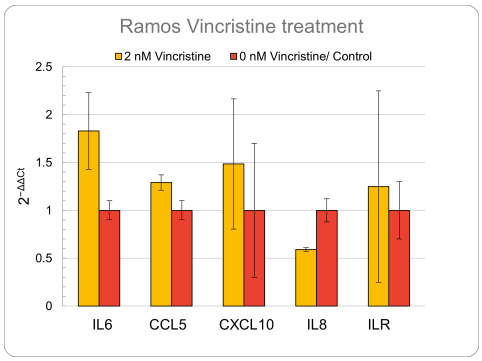
The qPCR analysis was performed on five human lymphoma cell lines (Ramos, Daudi, SU-DHL-10, and OCI-LY3) to assess their inflammatory response to Chromosomal Instability (CIN) induced by Vincristine treatment (0 nM, 2 nM, and 5 nM). The expression of IL6, CCL5, CXCL10, IL-8, and IL6R was evaluated.

IL6 expression was consistently upregulated across all viable cell lines following Vincristine treatment (Figure 6). SU-DHL-10 exhibited the highest fold-change in IL6 expression (12.16-fold at 2 nM Vincristine), followed by OCI-LY3 (3.78-fold at 2 nM), Daudi (2.39-fold at 5 nM and 1.58-fold at 2 nM), and Ramos (1.83-fold at 2 nM) (Figure 6).

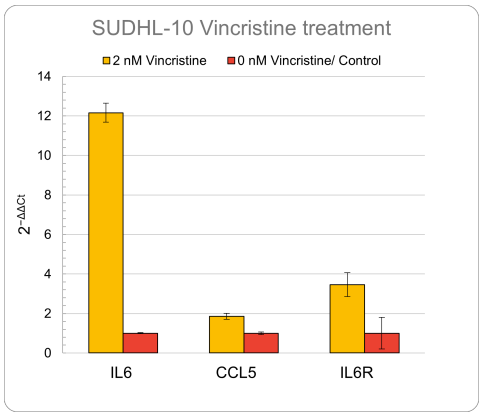
The Daudi cell line demonstrated the most consistent dose-dependent in-



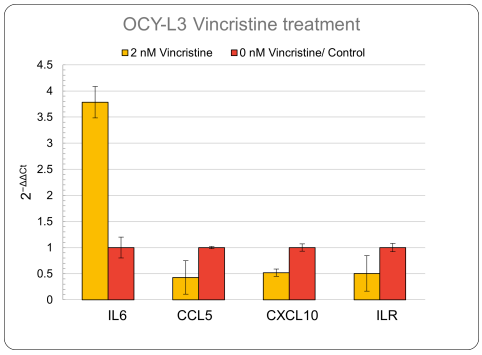
(A)



(B)



(C)



(D)

Figure 6. The Expression Fold Change of IL6, CCL5, CXCL10, IL8 and IL6-R. The double delta CT analysis of the qPCR results for cytokine expression of IL6, CCL5, CXCL10, IL8 and IL6-R and the technical standard deviation for each condition after treatment with 2 nM and 5 nM Vincristine, compared to the untreated control (0 nM) of the cell lines Daudi, Ramos, SU-DHL-10, and OCI-LY3. The data is presented as the mean fold change, also displaying the technical standard deviation as error bars. (A) Expression fold change of cytokines and IL6-R in Daudi. (B) Expression fold change of cytokines and IL6-R in Ramos. (C) Expression fold change of cytokines and IL6-R in SU-DHL-10. (D) Expression fold change of cytokines and IL6R in OCI-LY3.

crease in cytokine expression. At 5 nM Vincristine, IL6, CCL5, CXCL10, and IL-8 expression increased by 139%, 97%, 181%, and 191%, respectively, compared to the control. IL6R expression also increased by 32% (Figure 6A).

CCL5 expression varied across cell lines, showing dose-dependent increases in Daudi and moderate increases in Ramos and SU-DHL-10, but decreased in OCI-LY3. CXCL10 expression increased dose-dependently in Daudi, moderately in Ramos, but decreased in OCI-LY3 (Figure 6). IL-8 expression increased dose-dependently in Daudi but decreased in Ramos. IL6R expression increased in Ramos, Daudi, and SU-DHL-10 upon Vincristine treatment, but decreased in OCI-LY3 (Figure 6).

SU-DHL-6 was excluded from the analysis due to consistently high CT values (>35) for all tested genes. Similarly, cells treated with 5 nM Vincristine were excluded for all cell lines except Daudi due to high CT values. Subsequently, each cell line was statistically analyzed by applying a single factor ANOVA to test statistical significance.

In the Daudi cells F was shown to be $> F_{\text{critical}}$. Since $F = 6.13 > F_{\text{critical}} = 3.89$, the difference between groups is statistically significant, which is also underlined by the p-value of 0.0147. The Tukey's HSD test was then applied to identify which groups are different. The test showed no significance for the pairwise comparison of 5 nM Vincr versus 2 nM Vincr (HSD=0.834), and 2 nM Vincr versus Control (HSD=0.446). However, a statistical significance was shown for 5 nM Vincr versus Control (HSD=1.28).

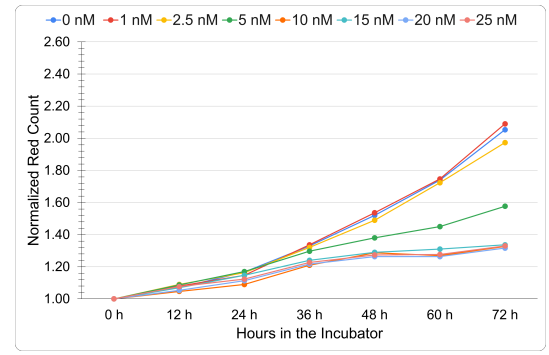
In the Ramos cells the p-value (0.192) is greater than the typical alpha level of 0.05, which means the difference between the treatment and control groups is not statistically significant. The F-statistic (2.031) is less than the F critical value (5.318), also indicating no significant difference.

In the SU-DHL-10 cells the p-value (0.206) is greater than 0.05 meaning the difference between the treatment and control groups is not statistically significant. The F-statistic (2.270) is less than the F critical value (7.709), also indicating no significant difference. While the Vincristine treatment has a much higher average value than the control (5.824 versus 1.000), the large variance in the treatment group (30.755) indicates there was substantial variability in the measurements. This high variability makes it difficult to conclude that the observed difference is due to the treatment. In the OCY-L3 cells the p-value (0.722) is much higher than 0.05, indicating that the difference between treatment and control groups is not statistically significant. The F-statistic (0.139) is substantially lower than the F critical value (5.987), also indicating no significant difference. The high variance in the treatment group (2.727) and the very high p-value suggest that this difference is likely due to random variation rather than a true treatment effect.

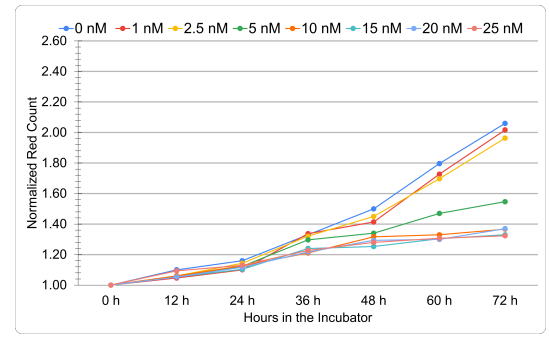
These results indicate that high dose Vincristine treatment induces an inflammatory response in the human lymphoma cell Daudi, characterized by the upregulation of IL-6 and variable changes in the expression of CCL5, CXCL10, IL-8, and IL-6R. The inflammatory response and statistical significance varied across different cell lines. However, the consistent upregulation of IL-6 across all cell lines suggests its crucial role in the cellular response to CIN-induction, potentially contributing to drug resistance or other adaptive mechanisms.

IncuCyte Analysis

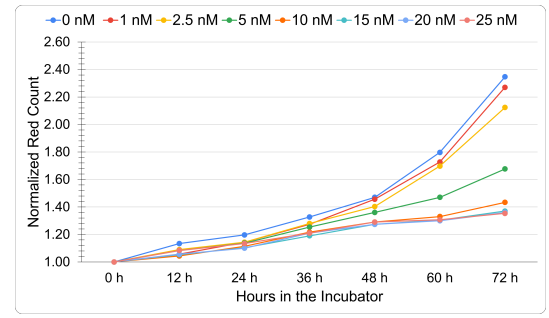
The study examined the dose-dependent effects of BAY1217389-induced CIN on cell growth and viability in human lymphoma cell lines using Incucyte analysis and Pearson correlation analysis. The research focused on three cell lines: Daudi, Ramos, and SU-DHL-10, each seeded at various densities and treated with BAY1217389 concentrations ranging from 0 to 25 nM over 72 hours.



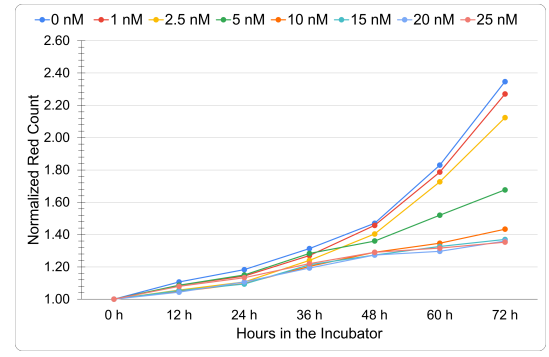
(A)



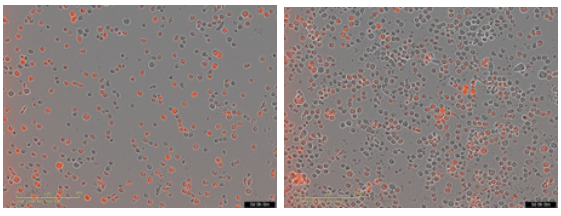
(B)



(C)



(D)



(E) (F)

Figure 7. Growth Rate of Daudi cells treated with 0 nM, 1 nM, 2.5 nM, 5 nM, 10 nM, 15 nM, 20 nM, and 25 nM BAY1217389. (A) Seeding density of 8000 cells. (B) Seeding density of 14000 cells. (C) Seeding density of 20000 cells. (D) Seeding density of 26000 cells. (E) IncuCyte cell image at $t = 0$ hours (left) versus $t = 72$ hours (right).

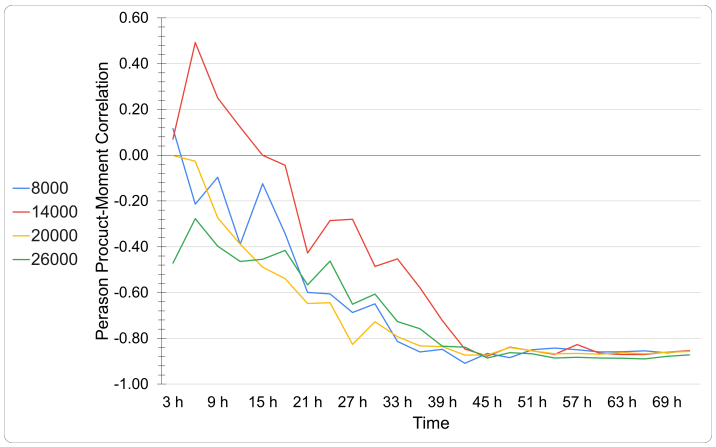


Figure 8. Pearson Product-Moment Correlation between the drug dose and normalized cell count. Quantification of the relationship between BAY1217389 concentration and cell prolifera-tion/viability over time in the Daudi cell line for each seeding density.

In the Daudi cell line higher concentrations of BAY1217389 (≥10 nM) consistently resulted in substantial inhibition of cell proliferation across all seeding densities. The anti-proliferative effects appeared more pronounced at higher cell densities (Figure 7). Pearson correlation analysis revealed strong negative correlations between drug concentration and normalized cell count, emerging rapidly at lower seeding densities and more gradually at higher densities (Figure 8).

The Ramos cell line showed dose-dependent growth inhibition at concentrations ≥5 nM across all seeding densities. Lower seeding densities exhibited more pronounced cytotoxic effects at higher drug concentrations (Figure 9). Pearson correlation analysis demonstrated initially weak correlations that became strongly negative over time, with the highest seeding density showing an initial positive correlation before turning negative (Figure 10).

For SU-DHL-10, concentrations ≥5 nM resulted in dose-dependent growth inhibition across both tested seeding densities. The lower seeding density (8,000 cells/well) showed more potent cytostatic or cytotoxic effects at higher concentrations (Figure 11). Pearson correlation analysis revealed rapid development of strong negative correlations at the lower seeding density, while the higher density initially showed strong positive correlations before shifting to negative values (Figure 12).

Across all cell lines, lower drug concentrations (1 nM and 2.5 nM) did not significantly impair cell proliferation compared to the control, suggesting a threshold concentration for anti-proliferative effects. The varying patterns of response across seeding densities indicate that initial cell density and potential cell-cell interactions may modulate the sensitivity and kinetics of the response to CIN induction.

These findings highlight the complex relationship between CIN induction, cell density, and proliferation in lymphoma cell lines, suggesting that the cellular response to CIN is influenced by both drug concentration and the initial cellular environment.

Discussion

This research aimed to investigate whether CIN in human lymphoma cells activates the IL6-mediated inflammatory signaling pathway and influences their viability, contributing to a pro-tumor microenvironment and modulating therapeutic outcomes. The findings from the genomics analysis, qPCR, and IncuCyte experiments collectively shed light on this research

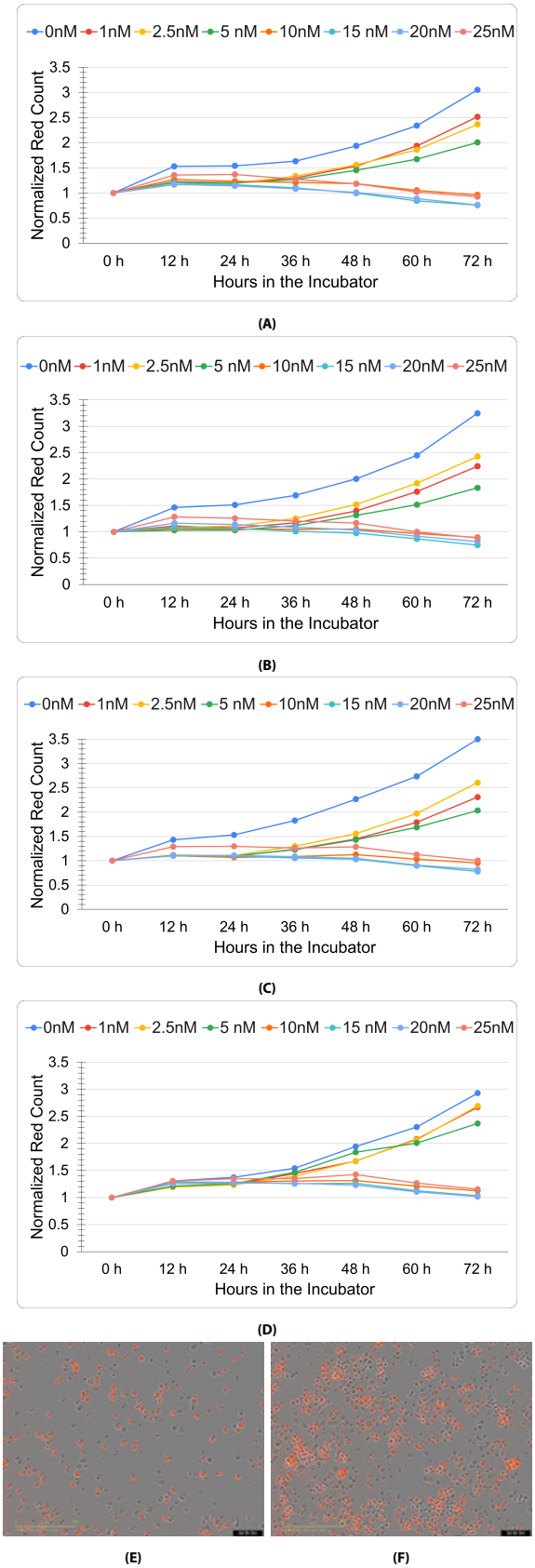


Figure 9. Growth Rate of Ramos cells treated with 0 nM, 1 nM, 2.5 nM, 5 nM, 10 nM, 15 nM, 20 nM, and 25 nM BAY1217389 (A) Seeding density of 8000 cells. (B) Seeding density of 14000 cells. (C) Seeding density of 20000 cells. (D) Seeding density of 26000 cells. (E) IncuCyte cell image at time point $t = 0$ hours (right) versus $t = 72$ hours (left).

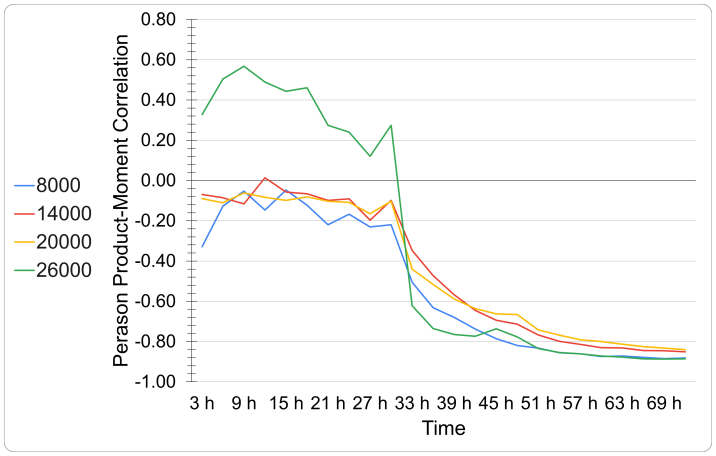


Figure 10. Pearson Product-Moment Correlation between the drug dose and normalized cell count. Quantification of the relationship between BAY1217389 concentration and cell proliferation/viability over time in the Ramos cell line for each seeding density.

question.

Chromosomal Instability

The genomics analysis revealed the diverse mutational landscape and genomic alterations present in the lymphoma cell lines, reflecting their varying degrees of genomic instability. This heterogeneity can promote the survival of subclones with altered inflammatory signaling pathways, as supported by the qPCR results. The OCI-LY3 cell line, with its high degree of genomic instability, demonstrated a substantial 3.78-fold elevation in IL6 expression upon 2 nM Vincristine treatment. This observation suggests that CIN could rewire inflammatory signaling pathways in lymphoma cells to promote tumorigenesis, aligning with findings from Cuceu et al. (2018) and Hong et al. (2022)[1,5].

Role of IL6 in CIN-Mediated Effects

Across all tested cell lines, the qPCR analysis revealed a consistent upregulation of IL6 expression upon Vincristine treatment, corroborating the findings of Hong et al. (2022)[5], who identified IL6 as a key regulator of the inflammatory response in chronic CIN. The SU-DHL-10 cell line, carrying concurrent rearrangements involving oncogenes (MYC and BCL2) and the IGH locus, exhibited the highest change (12.16-fold) in IL6 expression. This observation, coupled with the aggressive clinical course associated with these genetic alterations, suggests the potential role of IL6 in CIN-mediated lymphomagenesis.

While IL6 expression was consistently upregulated, the expression patterns of other cytokines and the IL6 receptor varied across cell lines. This heterogeneity in the inflammatory response, as seen in the qPCR results, could contribute to the differential effects observed in the IncuCyte analysis, where certain doses or cell lines exhibited unique kinetic patterns and responses to CIN induction.

The diverse expression patterns of other cytokines and the IL6 receptor (IL6R) may contribute to the differential effects observed in the IncuCyte analysis, where certain doses or cell lines exhibited unique kinetic patterns and responses to the CIN-inducing agent BAY1217389.

The findings from Hong et al. (2022)[5] provide complementary insights into the role of the IL6R in mediating CIN-induced effects. They observed a positive correlation between IL6R levels and the essentiality of genes involved in DNA repair, chromosome maintenance, and the mitotic spindle

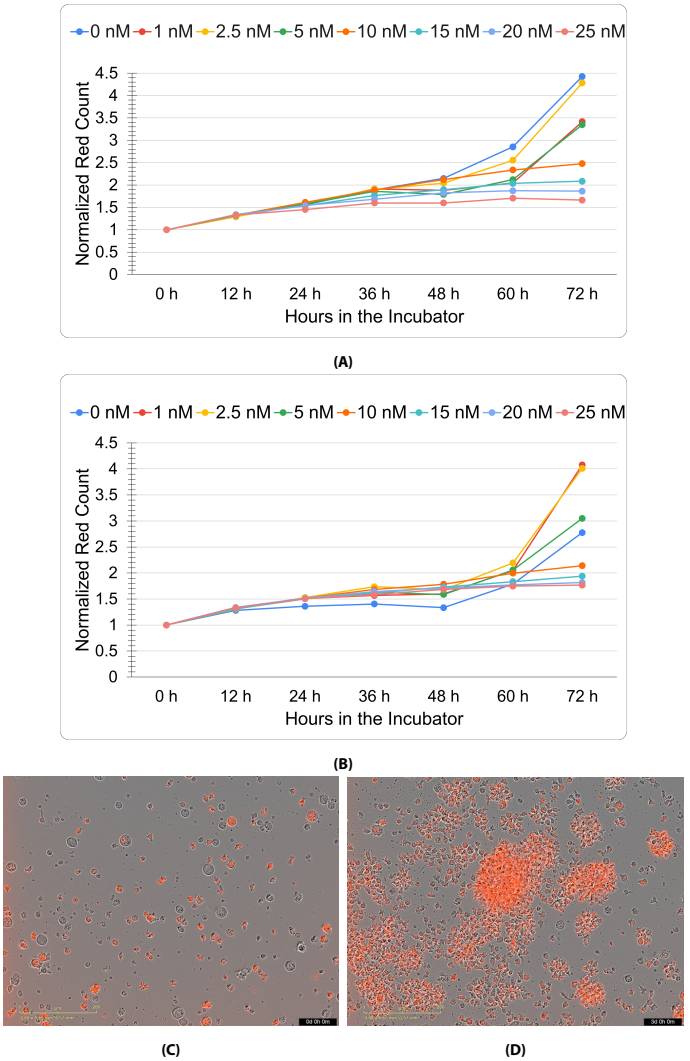


Figure 11. Growth Rate of SU-DHL-10 cells treated with 0 nM, 1 nM, 2.5 nM, 5 nM, 10 nM, 15 nM, 20 nM, and 25 nM BAY1217389. (A) Seeding density of 8000 cells. (B) Seeding density of 14000 cells. (C) IncuCyte cell image at timepoint $t = 72$ hours.

assembly checkpoint. This suggests that cancer cell lines with high IL6R levels are better equipped to regulate CIN across multiple cancer types. In the context of this study, the heterogeneous expression patterns of the IL6R across lymphoma cell lines may contribute to the observed differences in their responses to CIN induction.

Collectively, these findings support the hypothesis that inhibiting IL6 signaling, potentially in combination with CIN-inducing agents like Vincristine or Paclitaxel, could sensitize lymphoma cells to chemotherapy. The differential response to Vincristine treatment among cell lines with varying IL6 and IL6R expression levels further reinforces the notion that modulating the IL6 pathway may enhance therapeutic efficacy in lymphoma patients by targeting the CIN-induced inflammatory signaling mechanisms. However, it should be noted that a statistical significance was not found for all cell lines and all dosages of Vincristine making further research necessary.

Cell Viability After CIN Induction

The IncuCyte analysis revealed a dose-dependent effect of the CIN-inducing agent BAY1217389 on cell viability, with higher doses inducing more pronounced cytotoxic or cytostatic responses across the lymphoma cell lines. Interestingly, certain levels of CIN did not impair, and some

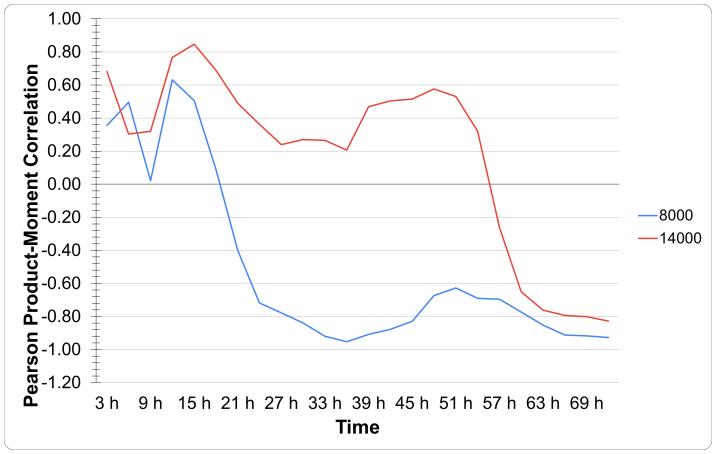


Figure 12. Pearson Product-Moment Correlation between the drug dose and normalized cell count. Quantification of the relationship between BAY1217389 concentration and cell proliferation/viability over time in the SU-DHL-10 cell line for each seeding density.

potentially even enhanced, cell growth in some cell lines, confirming the interplay between CIN and cell survival described by Cuceu et al. (2018)[1]. For instance, in the SU-DHL-10 cell line, lower drug concentrations supported cell growth, while higher concentrations inhibited growth in a dose-dependent manner (Figure 11).

The temporal dynamics and kinetic patterns observed in the IncuCyte analysis varied across doses and cell lines, indicating complex cellular responses to drug exposure and CIN induction. These findings, combined with the heterogeneity in the inflammatory response observed in the qPCR, suggest that CIN not only drives genetic diversity but also influences inflammatory signaling pathways that can modulate drug efficacy and therapeutic outcomes, as proposed by Turner & Reis-Filho (2012) and Hong et al. (2022)[5,14]. Additionally, the degree of cytotoxicity varied across different seeding densities, with lower seeding densities generally exhibiting greater susceptibility to the anti-proliferative and cytotoxic effects of BAY1217389. For instance, in the Ramos cell line, the lowest seeding density of 8,000 c/w showed the most potent cytotoxic effects at higher concentrations (15 nM, 20 nM, and 25 nM) compared to higher seeding densities (Figure 9). This observation suggests that lower cell densities may render lymphoma cells more vulnerable to the anti-proliferative and cytotoxic effects of CIN-inducing agents and that potential cell-cell interactions and cell-heterogeneity may modulate the sensitivity and kinetics of treatment responses.

Implications for Therapeutic Strategies

The consistent upregulation of IL6 expression across cell lines upon Vincristine treatment, as observed in the qPCR, coupled with the dose-dependent effects on cell viability and the differential responses observed in the IncuCyte analysis (Figure 9), support the hypothesis that CIN can activate the IL6-mediated inflammatory signaling pathway in lymphoma cells. The differential response to Vincristine treatment among cell lines with varying IL6 expression levels, as observed in the qPCR, suggests that IL6 may modulate therapeutic response in lymphoma. Combining IL6 inhibitors with standard chemotherapy regimens, such as Vincristine, may enhance therapeutic efficacy in lymphoma patients by targeting the CIN-induced inflammatory signaling pathways.

Limitations and Further Directions

Even though we provided insights into inflammatory signaling in human lymphoma cells, several limitations should be acknowledged. The limited

number of cell lines used in the qPCR analysis and the lack of data for some cytokines due to low cell viability after high-dose Vincristine treatment or double delta Ct values above 35, restricted the generalizability of the findings. The absence of a non-lymphoma cell line as a control group limited the ability to distinguish lymphoma-specific responses from more generalized cellular reactions to CIN-induced stress. The relatively low resolution of the IncuCyte analysis prevented the direct observation of the underlying mechanisms of CIN-induced changes in cell morphology and behavior, limiting the depth of mechanistic insights. As IL6 is a stress cytokine, its upregulation in response to CIN may not be specific to lymphoma cells, and similar responses have been observed in other cell lines, such as breast cancer cells[5]. For the IncuCyte analysis for the SU-DHL-10 cell line, there were not enough cells available after treatment with higher doses of BAY1217389, to test higher seeding densities (20,000 and 26,000 cells), limiting the data analysis. The CIN rates and specific genetic alterations in the cell lines need to be further characterized using techniques like single-cell DNA sequencing to better understand their influence on inflammatory signaling. Future research should focus on addressing these limitations by employing larger sample sizes, including control cell lines, and utilizing higher-resolution techniques to directly observe the effects of CIN on cell behavior and inflammatory signaling.

Conclusion

This research has provided valuable insights into the interplay between CIN, inflammatory signaling, and therapeutic responses in human lymphoma cells. The findings from the genomics analysis, qPCR, and IncuCyte experiments collectively suggested that CIN can activate the IL6-mediated inflammatory signaling pathway in lymphoma cells, contributing to a pro-tumor microenvironment and potentially modulating therapeutic outcomes. The consistent upregulation of IL6 expression observed across multiple lymphoma cell lines upon Vincristine treatment, coupled with the dose-dependent effects of BAY1217389 on cell viability, supports the hypothesis that CIN can influence inflammatory signaling and cellular responses to chemotherapeutic agents. The differential expression patterns of other cytokines and the IL6 receptor across cell lines further underscore the heterogeneity in inflammatory responses, which may contribute to the varying kinetic patterns and therapeutic responses observed in the IncuCyte analysis. By targeting the IL6 signaling axis in combination with standard chemotherapy regimens, it may be possible to enhance therapeutic efficacy and overcome treatment resistance in lymphoma patients. However, further research is necessary to optimize dosing and combinations, as well as to validate the findings in preclinical animal models and clinical studies.

Overall, this research contributes to the growing body of knowledge on the interplay between CIN, inflammatory signaling, and therapeutic responses in cancer, with a specific focus on lymphoma. The insights gained can contribute to the groundwork that might pave the way for the development of innovative therapeutic strategies that leverage the understanding of CIN-induced inflammatory signaling pathways to improve patient outcomes.

Acknowledgments

I would like to express my gratitude to the Laboratory of Genomic Instability in Development and Disease for providing the resources and facilities to conduct my research and experiments for this Paper. I would like to thank my supervisor, Dr. Floris Foijer, for his guidance and support throughout this journey. Moreover, I would also like to extend my thanks to Iris and Merlissa, who supervised me in the Laboratory and gave me valuable insights and explanations. Furthermore, I want to thank the entire team at the Laboratory of Genomic Instability in Development and Disease for creating a collaborative learning environment. Their willingness to answer my questions and share their knowledge was invaluable. Finally, I would like to acknowledge the University College Groningen, especially the science lecturers, for providing me with the foundational knowledge and academic support that laid the groundwork for this research.

References

1. Cuceu, C. et al. Chromosomal instability in Hodgkin lymphoma: an in-depth review and perspectives. *Cancers* **10,** 91 (2018). https://doi.org/10.3390/cancers10040091
2. Siri, S. O., Martino, J. & Gottifredi, V. Structural chromosome instability: types, origins, consequences, and therapeutic opportunities. *Cancers* **13,** 3056 (2021). https://doi.org/10.3390/cancers13123056
3. Kwon, M., Leibowitz, M. L. & Lee, J.-H. Small but mighty: the causes and consequences of micronucleus rupture. *Exp. Mol. Med.* **52,** 1777–1786 (2020). https://doi.org/10.1038/s12276-020-00529-z
4. Afonso, O., Figueiredo, A. C. & Maiato, H. Late mitotic functions of Aurora kinases. *Chromosoma* **126,** 93–103 (2017). https://doi.org/10.1007/s00412-016-0594-5
5. Hong, C. et al. cGAS–STING drives the IL-6-dependent survival of chromosomally instable cancers. *Nature* **607,** 366–373 (2022). https://doi.org/10.1038/s41586-022-04847-2
6. Kadowaki, H. & Nishitoh, H. Signaling pathways from the endoplasmic reticulum and their roles in disease. *Genes* **4,** 306–333 (2013). https://doi.org/10.3390/genes4030306
7. Chen, X., Shi, C., He, M., Xiong, S. & Xia, X. Endoplasmic reticulum stress: molecular mechanism and therapeutic targets. *Signal Transduct. Target. Ther.* **8,** 352 (2023). https://doi.org/10.1038/s41392-023-01570-w
8. Kwon, J. & Bakhoum, S. F. The cytosolic DNA-sensing cGAS–STING pathway in cancer. *Cancer Discov.* **10,** 26–39 (2020). https://doi.org/10.1158/2159-8290.CD-19-0761
9. Shi, Z. et al. Activation of the PERK-ATF4 pathway promotes chemo-resistance in colon cancer cells. *Sci. Rep.* **9,** 3210 (2019). https://doi.org/10.1038/s41598-019-39547-x
10. Wegiel, B., Bjartell, A., Culig, Z. & Persson, J. L. Interleukin-6 activates PI3K/Akt pathway and regulates cyclin A1 to promote prostate cancer cell survival. *Int. J. Cancer* **122,** 1521–1529 (2008). https://doi.org/10.1002/ijc.23261
11. Sempere, M. C., Fanjul, V. R., Pérez, I. S. & Perona, R. The role of the NFκB signalling pathway in cancer. *Clin. Transl. Oncol.* **10,** 143–147 (2008). https://doi.org/10.1007/s12094-008-0171-3
12. Johnson, D. E., O'Keefe, R. A. & Grandis, J. R. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat. Rev. Clin. Oncol.* **15,** 234–248 (2018). https://doi.org/10.1038/nrclinonc.2018.8
13. Braicu, C. et al. A comprehensive review on MAPK: a promising therapeutic target in cancer. *Cancers* **11,** 1618 (2019). https://doi.org/10.3390/cancers11101618
14. Turner, N. C. & Reis-Filho, J. S. Genetic heterogeneity and cancer drug resistance. *Lancet Oncol.* **13,** e178–e185 (2012). https://doi.org/10.1016/s1470-2045(11)70335-7

Submitted: 10/29/2024
Accepted: 03/06/2025
Published: 03/31/2025

Research Article

Simon Chen[1]

Improved Hardy-Sobolev Inequality under Moment Constraints

[1]Department of Physics, McGill University, Montréal, QC, Canada

Keywords

Geometric analysis, Closed manifolds, Sobolev inequality, Hardy-Sobolev inequality, Concentration-compactness principle

Email Correspondence

simon.chen2@mail.mcgill.ca

https://doi.org/10.26443/msurj.v20i1.253

Abstract

This paper is inspired by Aubin's 1979 result, which established that the best constant in the Sobolev inequality on the n-sphere, $\mathbb{S}^n$, can be improved under the condition of vanishing first-order moments. Recent advancements by Hang and Wang (2021) showed that Aubin's improvement can be generalized to arbitrary higher-order moments. We further extend Hang and Wang's results to the Hardy-Sobolev inequality on $\mathbb{S}^n$ by deriving an associated concentration-compactness principle and imposing similar moment constraints. Finally, we briefly outline a framework for extending these results to higher-order Sobolev spaces.

Introduction & Context

Sobolev Inequalities

Sobolev inequalities are fundamental tools in the study of mathematical analysis, geometry, and partial differential equations. They play a crucial role in embedding theorems, isoperimetric inequalities, and in ensuring the existence, uniqueness, and regularity of solutions to partial differential equations. Essentially, they establish a relationship between the L^p norms of functions and their derivatives, allowing one to trade regularity for integrability: a function that is sufficiently smooth (regularity) in L^p is also guaranteed to belong to a higher L^q space (integrability), where $q > p$.

To precisely state the classical Sobolev inequality, we first introduce the necessary function spaces. Let $n \geq 2$ and let $1 \leq p < n$. Denote by $\mathcal{D}(\mathbb{R}^n) = C_c^\infty(\mathbb{R}^n)$ the space of smooth functions with compact support and let $\mathcal{D}^{1,p}(\mathbb{R}^n)$ be its completion under the norm

$$\|u\| = \left(\int_{\mathbb{R}^n} |\nabla u|^p \right)^{1/p}. \tag{1}$$

This paper will primarily focus on functions in the Sobolev space $W^{1,p}$, but we will state Euclidean theorems in the larger space $\mathcal{D}^{1,p}$ for greater generality. Denote by $p^* = np/(n-p)$ the critical Sobolev exponent. The classical Sobolev inequality on $\mathbb{R}^n$, as proven by Sobolev [1, 2], is as follows:

Theorem 1 (Euclidean Sobolev Inequality)**.** *There exists a constant $C_{n,p} > 0$ such that, for any $u \in \mathcal{D}^{1,p}(\mathbb{R}^n)$,*

$$\left(\int_{\mathbb{R}^n} |u|^{p^*} \, dx \right)^{1/p^*} \leq C_{n,p} \left(\int_{\mathbb{R}^n} |\nabla u|^p \, dx \right)^{1/p}. \tag{2}$$

In the study of Sobolev inequalities, we are often interested in determining the smallest possible value of $C_{n,p}$ for which Theorem 1 remains valid. This minimal value for $C_{n,p}$ is often referred to as the best Sobolev constant. We will denote this best constant by $\mathbf{K}_{n,p}$. Rodemich [3], Aubin [4], and Talenti [5] proved that the best constant for $C_{n,p}$ exists and computed its value:

$$\mathbf{K}_{n,p} = \pi^{-1/2} n^{-1/p} \left(\frac{p-1}{n-p} \right)^{1-1/p} \left(\frac{\Gamma(n/2+1)\Gamma(n)}{\Gamma(n/p)\,\Gamma(n-n/p+1)} \right)^{1/n}.$$

Aubin [6] later extended the Euclidean Sobolev inequality (Theorem 1) to smooth, compact, Riemannian manifolds without boundary. The result is as follows:

Theorem 2 (Riemannian Sobolev Inequality)**.** *Let (M, g) be a smooth, closed, Riemannian n-manifold. Let $1 \leq p < n$ and let $p^* = np/(n-p)$. Then, for any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ that depends only on ε, M, and g such that, for any $u \in W^{1,p}(M)$,*

$$\left(\int_M |u|^{p^*} \, dv_g \right)^{p/p^*} \leq (\mathbf{K}_{n,p}^p + \varepsilon) \int_M |\nabla_g u|^p \, dv_g + C_\varepsilon \int_M |u|^p \, dv_g. \tag{3}$$

where ∇_g is the gradient with respect to the metric g and dv_g is the volume form on M.

Note that, for any Riemannian manifold, the Sobolev constant $\mathbf{K}_{n,p}^p + \varepsilon$ can be made arbitrarily close to $\mathbf{K}_{n,p}^p$ by choosing $\varepsilon > 0$ to be sufficiently small. However, unlike in the Euclidean case, the best Sobolev constant $\mathbf{K}_{n,p}^p$ cannot be achieved without the accompanying constant C_ε diverging.

Let $\mathbb{S}^n$ denote the n-sphere, the n-dimensional generalization of the 1-dimensional circle and the 2-dimensional sphere to any non-negative integer n. Aubin [7] showed that the Sobolev constant $\mathbf{K}_{n,p}^p + \varepsilon$ on $\mathbb{S}^n$ can be improved to $\mathbf{K}_{n,p}^p / 2^{p/n} + \varepsilon$ under the constraint that the first-order moments of $|u|^{p^*}$ vanish. In other words, if the function $|u|^{p^*}$ satisfies certain symmetry conditions, the value of the Sobolev constant $\mathbf{K}_{n,p}^p + \varepsilon$ can be lowered to $\mathbf{K}_{n,p}^p / 2^{p/n} + \varepsilon$. A precise statement of this result is as follows:

Theorem 3 (Aubin [7])**.** *Let $(\mathbb{S}^n, g_0)$ denote the n-sphere equipped with the round metric. Let $1 < p < n$ and let $p^* = np/(n-p)$. Then, for any*

$\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ that depends only on ε such that

$$\left(\int_{\mathbb{S}^n} |u|^{p^*}\, dv_{g_0}\right)^{p/p^*} \leq \left(\frac{\mathbf{K}_{n,p}^p}{2^{p/n}} + \varepsilon\right) \int_{\mathbb{S}^n} |\nabla u|^p\, dv_{g_0} + C_\varepsilon \int_{\mathbb{S}^n} |u|^p\, dv_{g_0}, \tag{4}$$

for any $u \in W^{1,p}(\mathbb{S}^n)$ that satisfies

$$\int_{\mathbb{S}^n} x_i |u|^{p^*}\, dv_{g_0} = 0 \tag{5}$$

for $i = 1, 2, ..., n+1$, where $(x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1}$.

Hang and Wang [8] further generalized Theorem 3 to higher-order moments. For consistency, we will follow their notation. In particular, denote by $\mathring{\mathcal{P}}_m$ the set of all polynomials $f : \mathbb{R}^{n+1} \to \mathbb{R}$ with degree at most m such that

$$\int_{\mathbb{S}^n} f\, dv_{g_0} = 0. \tag{6}$$

Additionally, for $0 < \theta < 1$ and $m \in \mathbb{N}$, define

$$\Theta(m, \theta, n) = \inf\Bigg\{ \sum_i \nu_i^\theta \;:\; \nu \text{ probability measure supported on countably many points } \{\xi_i\} \subseteq \mathbb{S}^n \text{ such that } \int_{\mathbb{S}^n} f\, d\nu = 0 \text{ for all } f \in \mathring{\mathcal{P}}_m, \nu_i := \nu(\{\xi_i\}) \Bigg\}. \tag{7}$$

Hang and Wang's [8] generalization of Aubin's improvement (Theorem 3) to higher-order moments is then as follows:

Theorem 4 (Hang and Wang [8])**.** *Let $(\mathbb{S}^n, g_0)$ denote the n-sphere equipped with the round metric. Let $1 < p < n$, let $m \in \mathbb{N}$, and let $p^* = np/(n-p)$. Then, for any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ that depends only on ε such that*

$$\left(\int_{\mathbb{S}^n} |u|^{p^*}\, dv_{g_0}\right)^{p/p^*} \leq \left(\frac{\mathbf{K}_{n,p}^p}{\Theta(m, p/p^*, n)} + \varepsilon\right) \int_{\mathbb{S}^n} |\nabla u|^p\, dv_{g_0} + C_\varepsilon \int_{\mathbb{S}^n} |u|^p\, dv_{g_0}, \tag{8}$$

for any $u \in W^{1,p}(\mathbb{S}^n)$ that satisfies

$$\int_{\mathbb{S}^n} f |u|^{p^*}\, dv_{g_0} = 0 \tag{9}$$

for all $f \in \mathring{\mathcal{P}}_m$.

Hang and Wang [8] also showed that $\Theta(1, p/p^*, n) = 2^{p/n}$ when $m = 1$, recovering Aubin's [7] original result (Theorem 3).

Hardy-Sobolev Inequalities

The Hardy-Sobolev inequality extends the Sobolev inequality to cases involving weighted integrals, where the integrand is multiplied by a weight function. In this paper, we will focus on weight functions of the form $d_g(x, x_0)^\alpha$, where $d_g(\cdot, x_0)$ denotes the Riemannian distance from a fixed point $x_0 \in M$. For the Euclidean case, this simplifies to $d_g(x, x_0)^\alpha = |x - x_0|^\alpha$, and we may set $x_0 = 0$ without loss of generality. Let $1 < p < n$, let $0 < \alpha < p$, and denote the critical Hardy-Sobolev exponent by $p_*(\alpha) = (n-\alpha)p/(n-p)$. Under these conditions, the Hardy-Sobolev inequality on $\mathbb{R}^n$ is as follows [9, 10]:

Theorem 5 (Euclidean Hardy-Sobolev Inequality)**.** *There exists a constant $C_{n,p,\alpha} > 0$ such that, for any $u \in \mathcal{D}^{1,p}(\mathbb{R}^n)$,*

$$\left(\int_{\mathbb{R}^n} \frac{|u|^{p^*(\alpha)}}{|x|^\alpha}\, dx\right)^{1/p^*(\alpha)} \leq C_{n,p,\alpha} \left(\int_{\mathbb{R}^n} |\nabla u|^p\, dx\right)^{1/p}. \tag{10}$$

Ghoussoub and Yuan [11] showed that the best constant for $C_{n,p,\alpha}$ can be attained. We denote this best constant by $\mathbf{K}_{n,p,\alpha}$. Egnell [12] computed its value, which is given by [13]

$$\mathbf{K}_{n,p,\alpha} = (n-\alpha)^{(n-p)/(n-\alpha)} \left(\frac{n-p}{p-1}\right)^{p-1} \times \left(\frac{\Gamma((p(n-\alpha)+p-n)/(p-\alpha))}{\Gamma(p(n-\alpha)/(p-\alpha))}\right)^{(p-\alpha)/(n-\alpha)} \times \left(\frac{n\pi^{n/2}\Gamma((n-\alpha)/(p-\alpha))}{\Gamma(n/2+1)}\right)^{(p-\alpha)/(n-\alpha)}. \tag{11}$$

Jaber [14], and Chen and Liu [15] later extended the Euclidean Hardy-Sobolev inequality (Theorem 5) to smooth, compact, Riemannian manifolds without boundary. We state their result:

Theorem 6 (Riemannian Hardy-Sobolev Inequality)**.** *Let (M, g) be a smooth, closed, Riemannian n-manifold and fix some $x_0 \in M$. Let $1 < p < n$, let $0 < \alpha < p$, and let $p^*(\alpha) = (n-\alpha)p/(n-p)$. Then, for any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ that depends only on ε, M, and g such that, for any $u \in W^{1,p}(M)$,*

$$\left(\int_M \frac{|u|^{p^*(\alpha)}}{d_g(x, x_0)^\alpha}\, dv_g\right)^{p/p^*(\alpha)} \leq (\mathbf{K}_{n,p,\alpha}^p + \varepsilon) \int_M |\nabla_g u|^p\, dv_g + C_\varepsilon \int_M |u|^p\, dv_g, \tag{12}$$

where d_g is the Riemannian distance on (M, g).

The goal of this paper is to improve the Hardy-Sobolev inequality (Theorem 6) on $\mathbb{S}^n$ by imposing moment constraints analogous to those used by Hang and Wang [8] to improve the standard Sobolev inequality.

Main Result

The following result presents an extension of Theorem 4 to the Hardy-Sobolev inequality:

Theorem 7. *Let $(\mathbb{S}^n, g_0)$ denote the n-sphere equipped with the round metric and fix some $x_0 \in \mathbb{S}^n$. Let $1 < p < n$, let $0 < \alpha < p$, let $m \in \mathbb{N}$, and let $p^*(\alpha) = (n-\alpha)p/(n-p)$. Then, for any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ that depends only on ε such that*

$$\left(\int_{\mathbb{S}^n} \frac{|u|^{p^*(\alpha)}}{d(x, x_0)^\alpha}\, dv_{g_0}\right)^{p/p^*(\alpha)} \leq \left(\frac{\mathbf{K}_{n,p,\alpha}^p}{\Theta(m, p/p^*(\alpha), n)} + \varepsilon\right) \times \int_{\mathbb{S}^n} |\nabla u|^p\, dv_{g_0} + C_\varepsilon \int_{\mathbb{S}^n} |u|^p\, dv_{g_0}, \tag{13}$$

for any $u \in W^{1,p}(\mathbb{S}^n)$ that satisfies

$$\int_{\mathbb{S}^n} f \frac{|u|^{p^*(\alpha)}}{d(x, x_0)^\alpha}\, dv_{g_0} = 0 \tag{14}$$

for all $f \in \mathring{\mathcal{P}}_m$.

To prove this theorem, we take inspiration from Hang and Wang [8] and instead prove the following, more general statement:

Theorem 8. *Let $(\mathbb{S}^n, g_0)$ denote the n-sphere equipped with the round metric and fix some $x_0 \in \mathbb{S}^n$. Let $1 < p < n$, let $0 < \alpha < p$, let $m \in \mathbb{N}$, let $p^*(\alpha) = (n-\alpha)p/(n-p)$, and let $T : \mathring{\mathcal{P}}_m \to \mathbb{R}_{>0}$ be some map. Then, for any $\varepsilon > 0$, there exists a constant $C_{\varepsilon,T} > 0$ that depends only on ε and T such that*

$$\left(\int_{\mathbb{S}^n} \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_{g_0}\right)^{p/p^*(\alpha)} \leq \left(\frac{\mathbf{K}^p_{n,p,\alpha}}{\Theta(m,p/p^*(\alpha),n)} + \varepsilon\right) \times \int_{\mathbb{S}^n} |\nabla u|^p\, dv_{g_0} + C_{\varepsilon,T}\int_{\mathbb{S}^n} |u|^p\, dv_{g_0}, \tag{15}$$

for any $u \in W^{1,p}(\mathbb{S}^n)$ that satisfies

$$\int_{\mathbb{S}^n} f\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_{g_0} \leq T(f)\left(\int_{\mathbb{S}^n} |u|^p\, dv_{g_0}\right)^{p^*(\alpha)/p} \tag{16}$$

for all $f \in \mathring{\mathcal{P}}_m$.

Since any $u \in W^{1,p}(\mathbb{S}^n)$ satisfying Equation 14 automatically satisfies Equation 16, Theorem 7 is recovered by taking the limit as $T(f) \to 0$.

A Concentration-Compactness Principle

The proof of Theorem 8 requires a concentration-compactness principle, similar to those introduced by Lions [10, 16]. The concentration-compactness principle is a key tool in the calculus of variations designed to address the lack of compactness in infinite-dimensional function spaces such as Sobolev spaces. Unlike in finite-dimensional spaces, where every bounded sequence has a convergent subsequence, sequences in infinite-dimensional spaces may fail to converge due to their mass either concentrating at specific points or escaping to infinity. The concentration-compactness principle provides a precise description of this phenomenon: any loss of compactness is restricted to a discrete, at most countable set of points. This weaker notion of compactness is often sufficient to analyze sequences that would otherwise be too irregular to handle. The classical concentration-compactness principle on $\mathbb{R}^n$ is given by Lemma I.1 in Lions [16]. A variant of this principle specifically associated to the Euclidean Hardy-Sobolev inequality is given by the following result:

Theorem 9 (Lemma 2.4 in Lions [10])**.** *Let $1 < p < n$ and let (u_k) be a bounded sequence in $\mathcal{D}^{1,p}(\mathbb{R}^n)$. Suppose that $u_k \to u$ pointwise almost everywhere, the sequence of measures $(|\nabla u_k|^p\, dx)$ is tight, and*

$$|\nabla u_k|^p\, dx \rightharpoonup \mu, \tag{17}$$

$$\frac{|u_k|^{p^*(\alpha)}}{|x|^\alpha}\, dx \rightharpoonup \nu, \tag{18}$$

where μ and ν are some measures on $\mathbb{R}^n$. Then there exists a non-negative real number $\nu_0 \in \mathbb{R}$ such that

$$\nu = \frac{|u|^{p^*(\alpha)}}{|x|^\alpha}\, dx + \nu_0\delta_0, \tag{19}$$

$$\mu \geq |\nabla u|^p\, dx + \mathbf{K}^{-p}_{n,p,\alpha}\nu_0^{p/p^*(\alpha)}\delta_0. \tag{20}$$

We require a generalization of Theorem 9 to smooth, closed, Riemannian manifolds. The result is as follows:

Theorem 10. *Let (M, g) be a smooth, closed, Riemannian n-manifold and fix some $x_0 \in M$. Let $1 < p < n$ and let (u_k) be a bounded sequence in $W^{1,p}(M)$. Suppose that $u_k \to u$ pointwise almost everywhere and that*

$$|\nabla_g u_k|^p\, dv_g \rightharpoonup \mu, \tag{21}$$

$$\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g \rightharpoonup \nu, \tag{22}$$

where μ and ν are some measures on M. Then there exists a non-negative real number $\nu_0 \in \mathbb{R}$ such that

$$\nu = \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g + \nu_0\delta_{x_0}, \tag{23}$$

$$\mu \geq |\nabla_g u|^p\, dv_g + \mathbf{K}^{-p}_{n,p,\alpha}\nu_0^{p/p^*(\alpha)}\delta_{x_0}. \tag{24}$$

Before we prove this concentration-compactness principle, we first state two lemmas. The first is a useful inequality, and the second is related to concentration-compactness on manifolds.

Lemma 1. *Let $x, y \in \mathbb{R}$, and let $a \geq 1$. Then,*

$$||x|^a - |y|^a| \leq a(|x|^{a-1} + |y|^{a-1})|x-y|. \tag{25}$$

Proof. The proof is trivial if either $x = 0$ or $y = 0$. Suppose then that $x \neq 0$ and $y \neq 0$. Consider the function $f(t) = t^a$. By the mean value theorem, there exists a $z \in \mathbb{R}$ between $|x|$ and $|y|$ such that

$$\begin{aligned} ||x|^a - |y|^a| &= |az^{a-1}(|x| - |y|)| \\ &= az^{a-1}||x| - |y||. \end{aligned}$$

Since the map $t \mapsto t^{a-1}$ is increasing for $a \geq 1$ and z lies between $|x|$ and $|y|$, it follows that

$$\begin{aligned} ||x|^a - |y|^a| &\leq a\max\{|x|^{a-1}, |y|^{a-1}\}||x| - |y|| \\ &\leq a(|x|^{a-1} + |y|^{a-1})||x| - |y||. \end{aligned}$$

By the reverse triangle inequality,

$$||x|^a - |y|^a| \leq a(|x|^{a-1} + |y|^{a-1})|x-y|.$$

■

Lemma 2 (Lions [16])**.** *Let (M, g) be a smooth, closed, Riemannian n-manifold. Let μ and ν be two bounded, non-negative measures on M and let $1 \leq p < q \leq \infty$. Suppose that there exists a constant $C_0 \geq 0$ such that, for any $\varphi \in C^\infty_c(M)$, the measures satisfy*

$$\left(\int_M |\varphi|^q\, d\nu\right)^{1/q} \leq C_0\left(\int_M |\varphi|^p\, d\mu\right)^{1/p}. \tag{26}$$

Then there exists an at-most countable set $\mathcal{I}$, a set of distinct points $\{x_i\}_{i\in\mathcal{I}}$ in M, and positive numbers $\{\nu_i\}_{i\in\mathcal{I}}$ such that

$$\nu = \sum_{i\in\mathcal{I}} \nu_i\delta_{x_i}, \tag{27}$$

$$\mu \geq C_0^{-p}\sum_{i\in\mathcal{I}} \nu_i^{p/q}\delta_{x_i}. \tag{28}$$

In particular,

$$\sum_{i\in\mathcal{I}} \nu_i^{p/q} < \infty. \tag{29}$$

Although Lions [16] proved the above lemma in $\mathbb{R}^n$, it can be easily adapted to manifolds since the proof in $\mathbb{R}^n$ does not make use of any properties unique to Euclidean space. We now prove Theorem 10:

Proof of Theorem 10. Let $v_k := u_k - u$. Then $v_k \to 0$ pointwise almost everywhere. Suppose also that

$$|\nabla_g v_k|^p \, dv_g \rightharpoonup \widetilde{\mu}, \tag{30}$$

$$\frac{|v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g \rightharpoonup \widetilde{\nu}, \tag{31}$$

where $\widetilde{\mu}$ and $\widetilde{\nu}$ are some bounded, non-negative measures on M.

We first prove the condition on ν. Let $\varphi \in C_c^\infty(M)$. Then $\varphi v_k \in W^{1,p}(M)$. By the Riemannian Hardy-Sobolev inequality (Theorem 6),

$$\left(\int_M \frac{|\varphi v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g\right)^{p/p^*(\alpha)} \leq (\mathbf{K}_{n,p,\alpha}^p + \varepsilon) \int_M |\nabla_g(\varphi v_k)|^p \, dv_g + C_\varepsilon \int_M |\varphi v_k|^p \, dv_g. \tag{32}$$

In the limit as $k \to \infty$, the left-hand side of Equation 32 goes to

$$\lim_{k\to\infty} \left(\int_M \frac{|\varphi v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g\right)^{p/p^*(\alpha)} = \left(\int_M |\varphi|^{p^*(\alpha)} \, d\widetilde{\nu}\right)^{p/p^*(\alpha)}.$$

We estimate the first term on the right-hand side of Equation 32 using Minkowski's inequality:

$$\begin{aligned}\left(\int_M |\nabla_g(\varphi v_k)|^p \, dv_g\right)^{1/p} &\leq \left(\int_M (|\nabla_g\varphi||v_k| + |\varphi||\nabla_g v_k|)^p \, dv_g\right)^{1/p} \\ &\leq \left(\int_M |\nabla_g\varphi|^p |v_k|^p \, dv_g\right)^{1/p} \\ &\quad + \left(\int_M |\varphi|^p |\nabla_g v_k|^p \, dv_g\right)^{1/p}.\end{aligned}$$

Then,

$$\begin{aligned}&\left|\left(\int_M |\nabla_g(\varphi v_k)|^p \, dv_g\right)^{1/p} - \left(\int_M |\varphi|^p |\nabla_g v_k|^p \, dv_g\right)^{1/p}\right| \\ &\quad \leq \left(\int_M |\nabla_g\varphi||v_k|^p \, dv_g\right)^{1/p}.\end{aligned}$$

Since φ is smooth and has compact support, its gradient is bounded on M by some constant $C > 0$,

$$\begin{aligned}&\left|\left(\int_M |\nabla_g(\varphi v_k)|^p \, dv_g\right)^{1/p} - \left(\int_M |\varphi|^p |\nabla_g v_k|^p \, dv_g\right)^{1/p}\right| \\ &\quad \leq C\left(\int_M |v_k|^p \, dv_g\right)^{1/p}.\end{aligned}$$

Since (v_k) is a bounded sequence in $W^{1,p}(M)$ and $v_k \to 0$ pointwise almost everywhere, then $v_k \rightharpoonup 0$ in $W^{1,p}(M)$. Since $1 < p < p^*$, it follows by the Rellich-Kondrachov theorem [17, 18] that $v_k \to 0$ in $L^p(M)$. Then, in the limit as $k \to \infty$, we have

$$\lim_{k\to\infty} \left|\left(\int_M |\nabla_g(\varphi v_k)|^p \, dv_g\right)^{1/p} - \left(\int_M |\varphi|^p |\nabla_g v_k|^p \, dv_g\right)^{1/p}\right| = 0.$$

Hence, the first term on the right-hand side of Equation 32 goes to

$$\begin{aligned}\lim_{k\to\infty} \int_M |\nabla_g(\varphi v_k)|^p \, dv_g &= \lim_{k\to\infty} \int_M |\varphi|^p |\nabla_g v_k|^p \, dv_g \\ &= \int_M |\varphi|^p \, d\widetilde{\mu}.\end{aligned}$$

Similarly, the second term on the right-hand side of Equation 32 goes to zero. Combining these results, we obtain

$$\left(\int_M |\varphi|^{p^*(\alpha)} \, d\widetilde{\nu}\right)^{1/p^*(\alpha)} \leq (\mathbf{K}_{n,p,\alpha}^p + \varepsilon)^{1/p} \left(\int_M |\varphi|^p \, d\widetilde{\mu}\right)^{1/p}.$$

Since this is true for all $\varepsilon > 0$,

$$\left(\int_M |\varphi|^{p^*(\alpha)} \, d\widetilde{\nu}\right)^{1/p^*(\alpha)} \leq \mathbf{K}_{n,p,\alpha} \left(\int_M |\varphi|^p \, d\widetilde{\mu}\right)^{1/p}. \tag{33}$$

It follows immediately from Lemma 2 that there exists an at-most countable set $\mathcal{I}$, a set of distinct points $\{x_i\}_{i\in\mathcal{I}}$ in M, and positive numbers $\{\widetilde{\nu}_i\}_{i\in\mathcal{I}}$ such that

$$\widetilde{\nu} = \sum_{i\in\mathcal{I}} \widetilde{\nu}_i \delta_{x_i}, \tag{34}$$

$$\widetilde{\mu} \geq \mathbf{K}_{n,p,\alpha}^{-p} \sum_{i\in\mathcal{I}} \widetilde{\nu}_i^{p/p^*(\alpha)} \delta_{x_i}. \tag{35}$$

And, in particular,

$$\sum_{i\in\mathcal{I}} \widetilde{\nu}_i^{p/p^*(\alpha)} < \infty. \tag{36}$$

Furthermore, let $\varphi \in C_c^\infty(M)$. By the triangle inequality,

$$\begin{aligned}&\left|\int_M \varphi \frac{|v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g - \int_M \varphi \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g + \int_M \varphi \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g\right| \\ &\quad \leq \int_M |\varphi| \left|\frac{|v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} - \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} + \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\right| \, dv_g.\end{aligned}$$

Since φ is smooth and has compact support, there exists a constant $C > 0$ such that

$$\begin{aligned}&\left|\int_M \varphi \frac{|v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g - \int_M \varphi \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g + \int_M \varphi \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g\right| \\ &\quad \leq C \int_M \left|\frac{|v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} - \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} + \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\right| \, dv_g.\end{aligned}$$

Consider the sequence (f_k) given by $f_k = |u_k|/d_g(x,x_0)^{\alpha/p^*(\alpha)}$. For simplicity, we also write $f = |u|/d_g(x,x_0)^{\alpha/p^*(\alpha)}$. Then,

$$\begin{aligned}&\left|\int_M \varphi \frac{|v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g - \int_M \varphi \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g + \int_M \varphi \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} \, dv_g\right| \\ &\quad \leq C \int_M \left||f_k - f|^{p^*(\alpha)} - |f_k|^{p^*(\alpha)} + |f|^{p^*(\alpha)}\right| \, dv_g.\end{aligned}$$

Since $u_k \to u$ pointwise almost everywhere, it follows immediately that $f_k \to f$ pointwise almost everywhere. Observe also that, by the Riemannian Hardy-Sobolev inequality (Theorem 6),

$$\begin{aligned}\left(\int_M |f_k|^{p^*(\alpha)}\, dv_g\right)^{p/p^*(\alpha)} &= \left(\int_M \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\right)^{p/p^*(\alpha)} \\ &\le (\mathbf{K}^p_{n,p,\alpha} + \varepsilon)\int_M |\nabla_g u_k|^p\, dv_g \\ &\quad + C_\varepsilon \int_M |u_k|^p\, dv_g.\end{aligned}$$

Since (u_k) is a bounded sequence in $W^{1,p}(M)$, we find that (f_k) must also be a bounded sequence in $L^{p^*(\alpha)}$. Then, by the Brézis-Lieb lemma [19],

$$\lim_{k\to\infty}\int_M \left||f_k - f|^{p^*(\alpha)} - |f_k|^{p^*(\alpha)} + |f|^{p^*(\alpha)}\right|\, dv_g = 0.$$

Hence, in the limit as $k \to \infty$, we have

$$\begin{aligned}\lim_{k\to\infty}\Bigg|\int_M \varphi\frac{|v_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g \\ + \int_M \varphi\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\Bigg| = 0.\end{aligned}$$

Equivalently,

$$\widetilde{\nu} = \nu - \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g.$$

Renaming each $\widetilde{\nu_i}$ to ν_i and rearranging the terms,

$$\nu = \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g + \sum_{i\in\mathcal{I}} \nu_i \delta_{x_i}. \tag{37}$$

It remains to show that the set $\mathcal{I}$ is a singleton. Let $\varphi \in C^\infty_c(M)$ be such that $\text{supp}(\varphi) \subseteq M\backslash\{x_0\}$. By the triangle inequality,

$$\begin{aligned}\left|\int_M \varphi\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\right| \\ \le \int_{\text{supp}(\varphi)} |\varphi|\left|\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} - \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\right| dv_g.\end{aligned}$$

Since φ is smooth and has compact support, there exists a constant $C > 0$ such that

$$\begin{aligned}\left|\int_M \varphi\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\right| \\ \le C\int_{\text{supp}(\varphi)} \left|\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha} - \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\right| dv_g.\end{aligned}$$

Since $\text{supp}(\varphi) \subseteq M\backslash\{x_0\}$ is compact and $d_g(x,x_0)$ is continuous and positive on $M\backslash\{x_0\}$, the function $1/d_g(x,x_0)^\alpha$ achieves a finite maximum on $\text{supp}(\varphi)$. We absorb this value into the constant C. Then,

$$\begin{aligned}\left|\int_M \varphi\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\right| \\ \le C\int_{\text{supp}(\varphi)} \left||u_k|^{p^*(\alpha)} - |u|^{p^*(\alpha)}\right| dv_g.\end{aligned}$$

By applying Lemma 1 and absorbing a factor of $p^*(\alpha)$ into the constant C, we obtain

$$\begin{aligned}\left|\int_M \varphi\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\right| \\ \le C\int_{\text{supp}(\varphi)} \left(|u_k|^{p^*(\alpha)-1} + |u|^{p^*(\alpha)-1}\right)|u_k - u|\, dv_g \\ \le C\int_M |u_k|^{p^*(\alpha)-1}|u_k - u|\, dv_g \\ + C\int_M |u|^{p^*(\alpha)-1}|u_k - u|\, dv_g.\end{aligned}$$

By Hölder's inequality with Hölder conjugates $p^*(\alpha)/(p^*(\alpha)-1)$ and $p^*(\alpha)$, we obtain

$$\begin{aligned}\left|\int_M \varphi\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\right| \\ \le C\left(\int_M |u_k|^{p^*(\alpha)}\, dv_g\right)^{(p^*(\alpha)-1)/p^*(\alpha)} \\ \times\left(\int_M |u_k - u|^{p^*(\alpha)}\, dv_g\right)^{1/p^*(\alpha)} \\ + C\left(\int_M |u|^{p^*(\alpha)}\, dv_g\right)^{(p^*(\alpha)-1)/p^*(\alpha)} \\ \times\left(\int_M |u_k - u|^{p^*(\alpha)}\, dv_g\right)^{1/p^*(\alpha)}.\end{aligned}$$

Since (u_k) is a bounded sequence in $W^{1,p}(M)$ and $1 < p^*(\alpha) < p^*$, it follows by the Rellich-Kondrachov theorem [17, 18], that (u_k) is also a bounded sequence in $L^{p^*(\alpha)}(M)$. We absorb this upper bound for $\|u_k\|^{p^*(\alpha)-1}_{p^*(\alpha)}$ into the constant C. Then,

$$\begin{aligned}\left|\int_M \varphi\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\right| \\ \le C\left(\int_M |u_k - u|^{p^*(\alpha)}\, dv_g\right)^{1/p^*(\alpha)} \\ + C\left(\int_M |u|^{p^*(\alpha)}\, dv_g\right)^{(p^*(\alpha)-1)/p^*(\alpha)} \\ \times\left(\int_M |u_k - u|^{p^*(\alpha)}\, dv_g\right)^{1/p^*(\alpha)}.\end{aligned}$$

Since (u_k) is a bounded sequence in $L^{p^*(\alpha)}(M)$ and $u_k \to u$ pointwise almost everywhere, we obtain by Fatou's lemma,

$$\begin{aligned}\int_M |u|^{p^*(\alpha)}\, dv_g &\le \liminf_{k\to\infty}\int_M |u_k|^{p^*(\alpha)}\, dv_g \\ &\le C^{p^*(\alpha)}.\end{aligned}$$

Absorbing this upper bound for $\|u\|^{p^*(\alpha)-1}_{p^*(\alpha)}$ into the constant C,

$$\begin{aligned}\left|\int_M \varphi\frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi\frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g\right| \\ \le C\left(\int_M |u_k - u|^{p^*(\alpha)}\, dv_g\right)^{1/p^*(\alpha)} \\ + C\left(\int_M |u_k - u|^{p^*(\alpha)}\, dv_g\right)^{1/p^*(\alpha)} \\ \le C\left(\int_M |u_k - u|^{p^*(\alpha)}\, dv_g\right)^{1/p^*(\alpha)}.\end{aligned}$$

Since (u_k) is a bounded sequence in $W^{1,p}(M)$ and $u_k \to u$ pointwise almost everywhere, then $u_k \rightharpoonup u$ in $W^{1,p}(M)$. Since $1 < p^*(\alpha) < p^*$, it follows by the Rellich-Kondrachov theorem [17, 18] that $u_k \to u$ in $L^{p^*(\alpha)}(M)$. Then, in the limit as $k \to \infty$, we have

$$\begin{aligned}\lim_{k\to\infty} \left| \int_M \varphi \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g - \int_M \varphi \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g \right| &\le \lim_{k\to\infty} C \left(\int_M |u_k - u|^{p^*(\alpha)}\, dv_g \right)^{1/p^*(\alpha)} \\ &= 0.\end{aligned}$$

Since this holds for all $\varphi \in C_c^\infty(M)$ satisfying $\mathrm{supp}(\varphi) \subseteq M \backslash \{x_0\}$, it must be that $\nu_i = 0$ for all $i \in \mathcal{I} \backslash \{0\}$. Hence, we may write $\mathcal{I} = \{0\}$ and

$$\nu = \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_g + \nu_0 \delta_{x_0}. \tag{38}$$

We now prove the conditions on μ. First, let $\varphi \in C_c^\infty(M)$. By the triangle inequality,

$$\begin{aligned}\left| \int_M \varphi |\nabla u_k|^p\, dv_g - \int_M \varphi |\nabla u|^p\, dv_g - \int_M \varphi |\nabla v_k|^p\, dv_g \right| \\ \le \int_M |\varphi|\, \big| |\nabla u_k|^p - |\nabla u|^p - |\nabla v_k|^p \big|\, dv_g.\end{aligned}$$

Since φ is smooth and has compact support, there exists a constant $C > 0$ such that

$$\begin{aligned}\left| \int_M \varphi |\nabla u_k|^p\, dv_g - \int_M \varphi |\nabla u|^p\, dv_g - \int_M \varphi |\nabla v_k|^p\, dv_g \right| &\le C \int_M \big| |\nabla u_k|^p - |\nabla u|^p - |\nabla v_k|^p \big|\, dv_g \\ &= C \int_M \big| |\nabla u_k|^p - |\nabla u|^p - |\nabla u_k - \nabla u|^p \big|\, dv_g.\end{aligned}$$

Since (u_k) is a bounded sequence in $W^{1,p}(M)$, it follows that (∇u_k) is a bounded sequence in L^p. Then, by the Brézis-Lieb lemma [19],

$$\lim_{k\to\infty} \int_M \big| |\nabla u_k|^p - |\nabla u|^p - |\nabla u_k - \nabla u|^p \big|\, dv_g = 0.$$

Hence, in the limit as $k \to \infty$, we have

$$\lim_{k\to\infty} \left| \int_M \varphi |\nabla u_k|^p\, dv_g - \int_M \varphi |\nabla u|^p\, dv_g - \int_M \varphi |\nabla v_k|^p\, dv_g \right| = 0.$$

Equivalently,

$$\begin{aligned}\mu &= |\nabla u|^p\, dv_g + \widetilde{\mu} \\ &\ge |\nabla u|^p\, dv_g + \mathbf{K}_{n,p,\alpha}^{-p} \sum_{i\in\mathcal{I}} \widetilde{\nu}_i^{p/p^*(\alpha)} \delta_{x_i} \\ &= |\nabla u|^p\, dv_g + \mathbf{K}_{n,p,\alpha}^{-p} \nu_0^{p/p^*(\alpha)} \delta_{x_0}.\end{aligned}$$

■

Proof of the Main Result

We now prove the main result, Theorem 8. For convenience, we restate it here:

Theorem. *Let $(\mathbb{S}^n, g_0)$ denote the n-sphere equipped with the round metric and fix some $x_0 \in \mathbb{S}^n$. Let $1 < p < n$, let $0 < \alpha < p$, let $m \in \mathbb{N}$, let $p^*(\alpha) = (n-\alpha)p/(n-p)$, and let $T : \mathring{\mathcal{P}}_m \to \mathbb{R}_{>0}$ be some map. Then, for any $\varepsilon > 0$, there exists a constant $C_{\varepsilon,T} > 0$ that depends only on ε and T such that*

$$\begin{aligned}\left(\int_{\mathbb{S}^n} \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_{g_0} \right)^{p/p^*(\alpha)} \le \left(\frac{\mathbf{K}_{n,p,\alpha}^p}{\Theta(m, p/p^*(\alpha), n)} + \varepsilon \right) \\ \times \int_{\mathbb{S}^n} |\nabla u|^p\, dv_{g_0} \\ + C_{\varepsilon,T} \int_{\mathbb{S}^n} |u|^p\, dv_{g_0},\end{aligned} \tag{39}$$

for any $u \in W^{1,p}(\mathbb{S}^n)$ that satisfies

$$\int_{\mathbb{S}^n} f \frac{|u|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_{g_0} \le T(f) \left(\int_{\mathbb{S}^n} |u|^p\, dv_{g_0} \right)^{p^*(\alpha)/p} \tag{40}$$

for all $f \in \mathring{\mathcal{P}}_m$.

Proof of Theorem 8. For simplicity, define

$$\beta := \frac{\mathbf{K}_{n,p,\alpha}^p}{\Theta(m, p/p^*(\alpha), n)} + \varepsilon. \tag{41}$$

Suppose, towards a contradiction, that Equation 15 does not hold. Then, for any $k \in \mathbb{N}$, there exists a sequence $(u_k) \subseteq W^{1,p}(\mathbb{S}^n)$ satisfying

$$\int_{\mathbb{S}^n} f \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_{g_0} \le T(f) \left(\int_{\mathbb{S}^n} |u_k|^p\, dv_{g_0} \right)^{p^*(\alpha)/p} \tag{42}$$

such that

$$\begin{aligned}\left(\int_{\mathbb{S}^n} \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_{g_0} \right)^{p/p^*(\alpha)} > \beta \int_{\mathbb{S}^n} |\nabla u_k|^p\, dv_{g_0} \\ + k \int_{\mathbb{S}^n} |u_k|^p\, dv_{g_0}.\end{aligned} \tag{43}$$

Since the left-hand side is finite by the Riemannian Hardy-Sobolev inequality, we may assume by rescaling that

$$\left(\int_{\mathbb{S}^n} \frac{|u_k|^{p^*(\alpha)}}{d_g(x,x_0)^\alpha}\, dv_{g_0} \right)^{p/p^*(\alpha)} = 1. \tag{44}$$

It follows immediately that

$$\int_{\mathbb{S}^n} |\nabla u_k|^p\, dv_{g_0} \le \frac{1}{\beta}, \tag{45}$$

$$\int_{\mathbb{S}^n} |u_k|^p\, dv_{g_0} \le \frac{1}{k}. \tag{46}$$

Since (u_k) and (∇u_k) are both bounded in $L^p(\mathbb{S}^n)$, it follows that the sequence (u_k) is also bounded in $W^{1,p}(\mathbb{S}^n)$, a reflexive Banach space. Then, by the Banach-Alaoglu theorem, the sequence (u_k) has a weakly convergent subsequence in $W^{1,p}(\mathbb{S}^n)$, which we denote as (u_{k_j}). Suppose that $u_{k_j} \rightharpoonup u$ weakly for some $u \in W^{1,p}(\mathbb{S}^n)$. Then $u_{k_j} \rightharpoonup u$ weakly in $L^p(\mathbb{S}^n)$ and $\nabla u_{k_j} \rightharpoonup \nabla u$ weakly in $L^p(\mathbb{S}^n)$. However, we also have from

Equation 46 that $u_k \to 0$ in $L^p(\mathbb{S}^n)$, and hence $u_k \rightharpoonup 0$ weakly in $L^p(\mathbb{S}^n)$. By the uniqueness of the weak limit, it follows that $u \equiv 0$. Since it is clear that $u_{k_j} \to 0$ in $L^p(\mathbb{S}^n)$ also, by the Riesz-Fischer theorem, we may pass to a further subsequence, which we still denote as (u_{k_j}), that converges pointwise almost everywhere to zero. Now, consider the following sequences of measures:

$$(\mu_j) := \left(|\nabla u_{k_j}|^p \, dv_{g_0}\right), \tag{47}$$

$$(\nu_j) := \left(\frac{|u_{k_j}|^{p^*(\alpha)}}{d_g(x, x_0)^\alpha} \, dv_{g_0}\right). \tag{48}$$

By Equations 44 and 45, these sequences satisfy $\sup_j \mu_j(K) < \infty$ and $\sup_j \nu_j(K) < \infty$ for each compact set $K \subseteq \mathbb{S}^n$. By the weak compactness for measures, there exist subsequences, which we still denote as (μ_j) and (ν_j), such that

$$\mu_j = |\nabla u_{k_j}|^p \, dv_{g_0} \rightharpoonup \mu, \tag{49}$$

$$\nu_j = \frac{|u_{k_j}|^{p^*(\alpha)}}{d_g(x, x_0)^\alpha} \, dv_{g_0} \rightharpoonup \nu, \tag{50}$$

where μ and ν are some bounded, non-negative measures on $\mathbb{S}^n$. Hence, by the concentration-compactness principle (Theorem 10) there exists a non-negative real number $\nu_0 \in \mathbb{R}$ such that

$$\nu = \nu_0 \delta_{x_0}, \tag{51}$$

$$\mu \geq \mathbf{K}_{n,p,\alpha}^{-p} \nu_0^{p/p^*(\alpha)} \delta_{x_0}. \tag{52}$$

Furthermore,

$$\begin{aligned}
\nu(\mathbb{S}^n) &= \int_{\mathbb{S}^n} d\nu \\
&= \lim_{k\to\infty} \int_{\mathbb{S}^n} \frac{|u_k|^{p^*(\alpha)}}{d_g(x, x_0)^\alpha} \, dv_{g_0} \\
&= 1.
\end{aligned}$$

And,

$$\begin{aligned}
\mu(\mathbb{S}^n) &= \int_{\mathbb{S}^n} d\mu \\
&= \lim_{k\to\infty} \int_{\mathbb{S}^n} |\nabla u_k|^p \, dv_{g_0} \\
&\leq \frac{1}{\beta}.
\end{aligned}$$

Then, for any $f \in \mathring{\mathcal{P}}_m$,

$$\begin{aligned}
\left|\int_{\mathbb{S}^n} f \, d\nu\right| &= \left|\lim_{k\to\infty} \int_{\mathbb{S}^n} f \frac{|u_k|^{p^*(\alpha)}}{d_g(x, x_0)^\alpha} \, dv_{g_0}\right| \\
&\leq \lim_{k\to\infty} T(f) \left(\int_{\mathbb{S}^n} |u_k|^p \, dv_{g_0}\right)^{p^*(\alpha)/p} \\
&\leq \lim_{k\to\infty} T(f) \left(\frac{1}{k}\right)^{p^*(\alpha)/p} \\
&= 0.
\end{aligned}$$

Therefore, ν satisfies the conditions presented in Equation 7. Then, by the definition of infimum,

$$\begin{aligned}
\Theta(m, p/p^*(\alpha), n) &\leq \nu_0^{p/p^*(\alpha)} \\
&\leq \mathbf{K}_{n,p,\alpha}^p \mu(\mathbb{S}^n) \\
&\leq \frac{\mathbf{K}_{n,p,\alpha}^p}{\beta}.
\end{aligned}$$

Hence,

$$\beta \leq \frac{\mathbf{K}_{n,p,\alpha}^p}{\Theta(m, p/p^*(\alpha), n)}. \tag{53}$$

This contradicts our definition of β as given in Equation 41. It must then be that Equation 15 holds. ■

Conclusion

In this paper, we established an improved Hardy-Sobolev inequality on $\mathbb{S}^n$ under moment constraints, extending the work of Hang and Wang [8] from the standard Sobolev setting to the Hardy-Sobolev setting. To achieve this, we first derived a concentration-compactness principle adapted to the Hardy-Sobolev inequality on smooth, compact, Riemannian manifolds without boundary. Our main result demonstrates that imposing moment constraints on functions in $W^{1,p}(\mathbb{S}^n)$ leads to a tighter upper bound on the Hardy-Sobolev constant, similar to the improvements obtained by Aubin [7] and Hang and Wang [8] for the Sobolev case.

Following the approach of Hang and Wang [8] in Section 4 of their paper, a natural direction for future work is to extend the above results to higher-order Sobolev spaces $W^{k,p}(M)$. We briefly sketch the framework. Aubin [4] previously extended the Euclidean higher-order Sobolev inequality to smooth, compact, Riemannian manifolds without boundary. This suggests that a similar extension should hold for the higher-order Hardy-Sobolev inequality. Such an extension would allow us to generalize the concentration-compactness principle stated in Theorem 10 from $W^{1,p}(M)$ to $W^{k,p}(M)$. These results would then yield a higher-order version of Theorem 8, establishing a Hardy-Sobolev equivalent to Theorems 4.1 and 4.2 in Hang and Wang's paper [8].

Acknowledgments

I would like to express my deepest gratitude to Professor Jérôme Vétois for his invaluable support and guidance throughout the summer. His patience, expertise, and insightful feedback have been instrumental in shaping the direction and outcome of this project. His mentorship throughout the years has also profoundly influenced my personal and academic growth. It has been a privilege to work under his supervision, and I am thankful for the time, effort, and care he devoted to my development.

I am also grateful to my dearest friend, Xuzhu (Ruth) Wang, for her unwavering support and encouragement. Her sharp observations brought me clarity whenever I hit a roadblock, and her presence has been a lasting source of strength and inspiration.

This research was supported by an undergraduate research award from the Natural Sciences and Engineering Research Council of Canada (NSERC), for which I am very thankful.

Editor's Note

The citation format was modified from the standard MSURJ format to prevent ambiguity with mathematical notation and content.

References

1. Sobolev, S. L. On a theorem of functional analysis. *Mat. Sb.* **4,** 471–498 (1938). https://doi.org/10.1090/trans2/034

2. Sobolev, S. L. *Applications of functional analysis in mathematical physics* (American Mathematical Society, 1963).

3. Rodemich, E. The Sobolev inequalities with best possible constants in *Analysis seminar at the California Institute of Technology*. 1–25 (1966).

4. Aubin, T. Espace de Sobolev sur les variétés riemanniennes [Sobolev space on Riemannian manifolds]. *Bull. Sci. Math.* **100,** 149–173 (1976).

5. Talenti, G. Best Constant in Sobolev Inequality. *Ann. Mat. Pura Appl.* **110,** 353–372 (1976). https://doi.org/10.1007/BF02418013

6. Aubin, T. Problèmes isopérimétriques et espaces de Sobolev [Isoperimetric problems and Sobolev spaces]. *J. Differ. Geom.* **11,** 573–598 (1976). https://doi.org/10.4310/jdg/1214433725

7. Aubin, T. Meilleures constantes dans le théorème d'inclusion de Sobolev et un théorème de Fredholm non linéaire pour la transformation conforme de la courbure scalaire [Best constants in the Sobolev embedding theorem and a nonlinear Fredholm theorem for the conformal transformation of scalar curvature]. *J. Funct. Anal.* **32,** 148–174 (1979). https://doi.org/10.1016/0022-1236(79)90052-1

8. Hang, F. & Wang, X. Improved Sobolev Inequality under Constraints. *Int. Math. Res. Not.* **2022,** 10822–10857 (2021). https://doi.org/10.1093/imrn/rnab067

9. Kohn, R., Caffarelli, L. & Nirenberg, L. First order interpolation inequalities with weights. *Compos. Math.* **53,** 259–275 (1984).

10. Lions, P. L. The Concentration-Compactness Principle in the Calculus of Variations. The Limit Case, Part 2. *Rev. Mat. Iberoam.* **1,** 45–121 (1985). https://doi.org/10.4171/rmi/12

11. Ghoussoub, N. & Yuan, C. Multiple solutions for quasi-linear PDEs involving the critical Sobolev and Hardy exponents. *Trans. Am. Math. Soc.* **352,** 5703–5743 (2000). https://doi.org/10.1090/S0002-9947-00-02560-5

12. Egnell, H. Elliptic Boundary Value Problems with Singular Coefficients and Critical Nonlinearities. *Indiana Univ. Math. J.* **38,** 235–251 (1989).

13. Angelo Alvino, V. F. & Trombetti, G. On the best constant in a Hardy–Sobolev inequality. *Appl. Anal.* **85,** 171–180 (2006). https://doi.org/10.1080/00036810500277405

14. Jaber, H. Hardy–Sobolev equations on compact Riemannian manifolds. *Nonlinear Anal. Theory Methods Appl.* **103,** 39–54 (2014). https://doi.org/10.1016/j.na.2014.02.011

15. Chen, N. & Liu, X. Hardy-Sobolev equation on compact Riemannian manifolds involving p-Laplacian. *J. Math. Anal. Appl.* **487,** 123992 (2020). https://doi.org/10.1016/j.jmaa.2020.123992

16. Lions, P. L. The Concentration-Compactness Principle in the Calculus of Variations. The Limit Case, Part 1. *Rev. Mat. Iberoam.* **1,** 145–201 (1985). https://doi.org/10.4171/rmi/6

17. Rellich, F. Ein Satz über mittlere Konvergenz [A theorem on mean convergence]. *Nachr. Ges. Wiss. Gött., Math.-Phys. Kl.* **1930,** 30–35 (1930).

18. Kondrachov, W. Sur certaines propriétés des fonctions dans l'espace [On certain properties of functions in space]. *C. R. (Doklady) Acad. Sci. URSS (Nouv. Sér.)* **48,** 535–538 (1945).

19. Brézis, H. & Lieb, E. A relation between pointwise convergence of functions and convergence of functionals. *Proc. Am. Math. Soc.* **88,** 486–490 (1983). https://doi.org/10.1090/S0002-9939-1983-0699419-3

Submitted: 01/14/2024
Accepted: 03/05/2025
Published: 03/31/2025

Research Article

[1]Department of Biology, McGill University, Montréal, QC, Canada

Keywords

Blue mussel, Beech, Dynamic energy budget, Biogenic habitat, Facilitative habitat

Email Correspondence

brian.schatteman@mail.mcgill.ca

https://doi.org/10.26443/msurj.v20i1.216

Brian Schatteman[1]

Positive Niche Construction: Incorporating Facilitative Microhabitat into Mechanistic Niche Modeling

Abstract

Researchers in the fields of conservation biology and invasion ecology aim to predict the dispersal of species in a reproducible manner, based on quantifiable relationships between target organisms and their environment, through a process known as mechanistic niche modelling. By identifying physiological constraints unique to an organism and calculating its budget of key resources in a given location, sophisticated estimates of potential activity and fecundity can be developed. These spatially-dependent dynamic energy budgets (DEBs) currently ignore the general phenomenon of biological facilitation and, in particular, positive niche construction, wherein a single species improves the suitability of their local habitat for future individuals by modifying their own microclimate. From bed-forming mussels to canopy-forming trees, diverse species can modify their own microclimates by increasing habitat complexity and, in doing so, ameliorate the same physical stressors explicitly considered in mechanistic niche models. To demonstrate the applicability of facilitative habitat in these models, this study selects two habitat-forming organisms and employs DEB-based hindcasting tools to simulate (1) the growth and allocation impacts of temperature regulation in *Mytilus edulis* (blue mussel) beds and (2) the near-ground micrometeorological impacts of *Fagus* (beech) tree canopy coverage. A significant reduction in growth and reproductive capacity in blue mussels beyond their optimal temperature and an overall amelioration of temperature and water stress below the beech canopy were observed. Although further research is required to refine the microclimate and micrometeorological impacts assumed for these model organisms, these results suggest that maturing around conspecifics can facilitate persistence in otherwise poor quality habitat. Thus, both species are predicted to have a significant Allee niche, demonstrating the need to incorporate facilitative habitat into mechanistic niche models, especially those used to predict climate change-induced range shifts.

Introduction

Part 1: Microhabitat Regulation by Blue Mussels

Because vegetative growth increases the vertical and horizontal structural complexity of landscapes, nearly all vegetation-dominated habitats (forests, marshes, etc.) can be considered biogenic[1]. In marine ecosystems, these facilitated environments can be critical for the growth of diverse species, e.g. juvenile fish preferentially inhabit bryozoan beds[2]. Biogenic habitats often harbor complex ecosystems characterized by high biodiversity[3], in part due to organisms experiencing reduced physical stress through the physically-ameliorating traits of structure-forming species[1], whose facilitation of other species is often a byproduct of their normal activity[2]. For example, canopy trees in dense forests insulate the subcanopy from temperature extremes and increase subcanopy humidity by evapotranspiring and retaining water in their own biomass and in the soil below[1]. Even small forest plants, like subcanopy shrubs and bracken, can reduce summertime soil temperatures and increase water availability[1]. Meanwhile, in intertidal zones, aggregated seaweeds and invertebrates form beds that protect species below this "canopy" from heat and desiccation[1]. Beds of sessile mussels can reduce peak temperatures in their immediate microclimate by 20 °C. This is a functional trait that can combat the thermal stress characteristic of non-optimal climate latitudes, leading to unexpected persistence patterns for resident species[1]. Mussels themselves also alter water clarity and sediment chemistry by filter feeding[2], further influencing the environment and thus evolutionary pressures experienced by local species.

Microclimate modifications like this can complicate assessments of invasion potential. Along the South African intertidal, for example, the indigenous mussel and ectotherm *Perna perna* periodically opens its shell to aerobically respire[4]. While this behavior results in no body temperature effects in solitary mussels, in aggregate beds of conspecifics, researchers have observed evaporative cooling, leading to decreased group body temperatures and thus reduced desiccation and heat stress[4]. The invasive mussel *Mytilus galloprovincialis* does not exhibit this behavior but has managed to successfully invade the southern coast of South Africa, despite desiccation pressures[4]. After researchers manipulated the densities of native *Perna perna* and invasive *M. galloprovincialis* mussels in intertidal beds along this coast, they found that both the native and invasive mussel benefitted from group evaporative cooling facilitated by *Perna perna*[4]. Thus, the persistence of *M. galloprovincialis* outside its native range may be enabled by the thermoregulating presence of *Perna perna*, highlighting the need to consider stress-ameliorated habitat in predicting the limits of invasion by *M. galloprovincialis*[4].

Although obligate associations, i.e., interspecific relationships required for the persistence of at least one partner between habitat-forming species and resident heterospecifics (other species), are more common in tropical systems[3], these associations remain ecologically important in Canada and support the productivity of many fisheries[2]. Marine mussels like *Mytilus edulis* (blue mussel) can provide complex structure to marine substrate as a single species by producing byssus threads. Assemblages of organisms can also build biogenic habitat[2]. Thus, understanding the natural population

dynamics of species residing in facilitative habitat and assessing the net impact of human activities on this regime requires knowledge of numerous supporting species[2].

Loss of quality habitat caused by anthropogenic climate change disproportionately threatens freshwater and coastal organisms and compounds other anthropogenic stressors, leading to accelerated biodiversity loss[2]. For example, climate change exacerbates the heat and desiccation stresses characteristic of the rocky intertidal[1]. However, one investigation of warming impacts on mussel bed inhabitants in the Bodega Marine Reserve observed a strong decoupling of microhabitat conditions from external climate variation. The study found no significant relation between simulated hot-dry climate change scenario and frequency of lethal desiccation in isopods and crabs, despite evidence for mortality in these species being closely associated with co-occurrent desiccation and thermal stress outside this buffered habitat[1]. The reliance of these species on mussel beds poses an indirect, unique vulnerability to climate change, however, as biogenic habitats often exhibit very slow recovery times[3]. The stress response of these taxa may therefore be characterized by a tipping point, up until which the mortality of buffered species may be unaffected until the engineering species itself is sufficiently stressed to cause the breakdown of the facilitative habitat and with it, the facilitated community[1].

The strength and prevalence of microclimate manipulation abilities in structure-forming species suggests that estimations of current and future distributions of these species and local residents must consider the hysteresis of habitat suitability, as potentially lethal macroclimatic conditions may conceal the presence of a nonlethal microclimate[1]. The uncertainty and potential lethality of climate change for myriad species and biogenic habitats themselves justifies studying microclimate-modifying species and the dynamics of physical facilitation[1]. Likewise, the recent acceleration of global biological invasions demands better predictions of the invasion potential for these habitat-forming species and their facilitated heterospecifics[1].

To estimate current habitat suitability and predict future range shifts, ecologists have turned to mechanistic niche modelling. This class of methods relies on ecophysiological data, such as target species-specific minimum/maximum temperature tolerances. They reduce the global landscape of climatic variation down into patches where the resource needs and environmental constraints of the target species are satisfied, thus identifying its potential distribution[5]. While these models generate more explanatory predictions of persistence than their correlative counterparts, they intrinsically neglect some aspects of the target species' biology, including dispersal ability, behavior, competitive fitness, and genetics. Additionally, many target species effectively persist in facilitative habitats, meaning that mechanistic niche models based on individual species' traits will fail to recognize the spatial boundaries of all quality biogenic habitat[1]. Their strict focus on the physiology and behavior of a target species overestimates the importance of these parameters, as the population dynamics and dispersal of a given species may be altogether controlled by a facilitating, microclimate-modifying counterpart[1].

Mechanistic niche models can employ Dynamic Energy Budgets (DEBs) to estimate the ability of a species to persist within a physiologically relevant set of climate conditions by calculating the net resource surplus accessible to the organism given its allocation needs, activity costs, and resource availability[5]. Building a dynamic energy budget for a target species requires estimating various metrics, called DEB parameters, of that organism's allocation strategy, reproductive costs, respiration costs, nutritional needs, and climate tolerances. To expand the biological traits considered in traditional DEB-based mechanistic niche modelling and to quantify the impact of conspecific habitat facilitation, also called positive niche construction, I here select *Mytilus edulis* (blue mussel) as a target species known to facilitate microhabitat by reducing local thermal stress[6]. I then compare its predicted persistence ability with and without the ameliorating influence of a conspecific mussel bed at a specific site in its native range.

Part 2: Micrometeorology Below Beech Trees

Key metrics of microclimate, including solar radiation, air and soil temperature, water availability, and wind, are significantly modified by forest vegetation. This leads to distinct microclimate profiles along vegetation type and land use axes[7]. Simultaneously, local plants adapt to ambient microclimatic conditions, e.g. air temperature controls respiration, water transport, and soil formation rates[8]. This simultaneous feedback between forest plants and their immediate microclimate creates an ecological system where the dynamics of individual elements are mutually dependent[7].

The heterogeneous microclimate landscape within forests leads to characteristically high subcanopy biodiversity, as the competitive fitness of individual plant species changes with the environmental conditions they experience at an organismal scale[8]. For example, light availability alone influences seed germination, recruitment, and establishment[9]. Forest plants also respond to microclimatic variation physiologically, e.g. beech trees exhibit reduced water uptake, root growth, foliage mass, and radial growth with lower wintertime soil temperatures[8]. Across taxa, branches adapt architecturally to light availability, and the inclination angle of leaves themselves reflects available solar radiation[7]. Throughout maturation, the relative abundance of sun and shade optimized leaves also changes depending on light availability[7].

Concurrently, forest plants, especially canopy trees, significantly modify the microclimate experienced by other forest residents[10]. Below the canopy, species experience lower maximum temperatures, higher minimum temperatures, and higher humidity[11]. This suggests that large trees moderate annual and seasonal climatic variation below the canopy, leading to a habitat with ameliorated physical stress conditions[11]. Overstory foliage further regulates the transmittance of light–that is, the amount that reaches the understory[9]. Canopy composition can also affect this quantity, as species-specific porosity, height, and spacing all impact the transmittance, quality, and variability of understory light[9,11]. Structural properties of these canopy species, such as leaf area index (LAI) and leaf distribution[7], can be used to calculate light transmittance according to Beer-Lambert's Law[9,11]. All forest management activities that adjust crown structure, canopy openness, and community composition thereby affect the amount of solar radiation available in the subcanopy[7,11]. This microclimate parameter regulates understory plant diversity, productivity, and competitive dynamics[12], and thus, controls on this parameter have ecosystem-wide consequences.

Overall, forest vegetation reduces local wind speed, depending on the size and distribution of plant biomass[7]. However, gaps in the canopy can generate turbulence and increase the prevalence of small turbulent eddies[7]. This explains the intentional placement of meteorological weather stations in wide clearings to better represent regional rather than localized conditions[13]. Furthermore, because of high evapotranspiration rates in forests, relative humidity is greater below forest canopies than above clearings[7].

Although the temperature profile within a forest varies throughout the day and year, in general, the presence of forest species reduces daily and seasonal temperature variation when compared to open ground, i.e. soils below a developed canopy exhibit warmer winter temperatures and colder summer temperatures than their exposed counterparts[7]. Soil temperature has remained an understudied aspect of microclimate despite being shown to significantly affect the photosynthesis, respiration, and growth of forest species[7]. However, in forestry, the importance of this sheltering phenomenon in stand growth and recruitment has already been established[7]. Microclimate impacts of dense forest cover can even be seen in adjacent disturbed patches, in a phenomenon known as "forest influence"[14]. The reduced dispersal distance and physical amelioration of this adjacency effect

can lead to faster recolonization of disturbed patches by mature forest communities[14]. This explains the increasingly common silvicultural practice of retention forestry, wherein mature patches are left within harvested areas to accelerate regeneration and, incidentally, recover biodiversity[14].

While these sheltered environments ameliorate some climatic stresses like drought and wind throw, the net impact of forest cover on recruitment is not linear, as many seedlings respond best to partial thinning of forest cover achieved via disturbance or harvesting[14]. In silviculture, juvenile growth appears optimized at canopy coverages between 25% and 75% depending on the ecophysiology of the cultivar[7]. This growth preference likely stems from the combined effects of reduced resource competition and remnant forest influence. The nontrivial dynamics of forest influence demonstrate the complexity of assessing growth conditions at a given stand density.

Despite the uncertainty regarding subcanopy light availability and thus growth suitability, the positive impact of forestation on temperature and humidity demonstrates the ability of forest cover to insulate the subcanopy environment from macroclimatic variation[15]. At local scales, this temperature buffering capacity depends on the thermodynamic efficiency of the ecosystem, as forest stands with a more homogenous distribution of biomass (e.g., a plantation) absorb and dissipate solar radiation more efficiently than stands structured heterogeneously (e.g., a mature, naturally regenerating forest)[11]. Accordingly, the amount of litter in a forest habitat affects its thermodynamic efficiency, soil evaporation rate, water retention capacity, and thus macroclimate buffering capacity[11]. It follows that to build accurate heat and dynamic energy budgets for resident species, researchers must study the thermodynamic efficiency of forest ecosystems[11]. Furthermore, this quantity reflects the ability of an ecosystem to buffer broader scale global warming and with it, thermophilization, or phase shifts towards species better adapted to higher temperatures[11].

One assessment of plant community thermophilization relied on correlative niche modelling to calculate the history of floristic temperature at resurveyed vegetation plots in North America and Europe based on their unique community assemblages[15]. The authors observed a higher frequency of warm-adapted understory plant species with time. This thermophilization occurred more rapidly in areas with higher warming rates[15]. In forests that became denser over the study period, thermophilization occurred more slowly, suggesting that the closure of temperate forests has historically insulated understory plant communities from macroclimatic warming and slowed associated phase shifts[15]. In another study that compared the microclimate-buffering capacities of primary and secondary forests, plantation forests exhibited < 2.5 °C hotter understories than their old growth counterparts, which were characterized by higher biomass throughout various vegetation levels[13]. Overall, this significantly greater ability of old-growth forests to buffer macroclimatic change, even at similar canopy densities, demonstrates the need to critically examine management practices[13]. Such examination is essential if stakeholders are to leverage habitat facilitation to slow biodiversity loss in complex ecosystems.

Forests' ability to ameliorate subcanopy stress and thus facilitate viable microhabitat invites the challenge of incorporating this facilitation into contemporary mechanistic niche models to improve our estimates of current and future biogeography. One way to acknowledge this phenomenon is to estimate the temperature impacts of a particular canopy species on the understory and identify changes to the suitability of that facilitated microhabitat. In light of its disproportionately severe reduction of near-ground solar radiation at a given canopy coverage, this study uses the genus *Fagus* (beech) to analyze the near-ground micrometeorological effects of tree coverage[11]. It has been shown that subcanopy solar radiation significantly affects beech recruitment by affecting LAI growth, mainly during early growth stages[12]. Thus, by comparing the influences of canopy coverage and macroclimatic warming on the microclimate experienced by juvenile beech trees, the importance of positive niche construction to this canopy species can be estimated.

Methods

Part 1: Microhabitat Regulation by Blue Mussels

To simulate the effect of increased body temperature on various life history, allocation, and growth parameters in blue mussels, I used the Dynamic Energy Budget Model in the Sea (Kearney & Porter[5]), one of the Biological Forecasting and Hindcasting Tools developed by Professor Michael Kearney's group at the University of Melbourne. This tool relies on the DEB modelling function built into NicheMapR, an R package developed by that same group to simulate microclimate conditions and thus produce mechanistic niche models of endothermic and exothermic organisms[5]. Calculations of blue mussel ecophysiology (growth, weight, reproduction, oxygen consumption, etc.) under modeled conditions thus derive from a mass budget scheme which can be described by stoichiometric equations of the standard model of Dynamic Energy Budget theory (see Supplementary Figure 5). This function is informed by DEB parameters sourced from the Add-my-Pet (AmP) collection, a self-described open-access scientific journal to which researchers contribute their findings on animal species-specific energetics, including DEB parameters, and their methods for deriving these quantities and associated species traits[16].

The DEB Model in the Sea hindcasting tool utilizes sea surface temperature data provided by the National Oceanic and Atmospheric Administration (NOAA) (Kearney & Porter[5]). Species-specific DEB parameters (z – size multiplier, κ – allocation fraction of growth and somatic maintenance, and Tb – body temperature) are input by the user with the model conditions (initial stage, ageing, post-hatch respiration, location). Having collated the required DEB parameters from various sources[16–21] (z, κ, and Tb for the DEB Model in the Sea, see Supplementary Figure 4), I analyzed the (1) total wet mass gain, (2) change in reproduction buffer, and (3) onset of puberty in maturing blue mussels initially at the egg stage. The 365-day simulations occurred from January 31, 2020, to January 31, 2021, at coordinates 60° N, 1° E. Mussels were subject to initial body temperatures of 18, 20, 22, 24, 26, 28, and 30 °C (see Supplementary Figure 1 for an example of parameter input). The output allocation and growth graphs were then visually analyzed for puberty onset, reproductive buffer, and wet mass gain.

Part 2: Micrometeorology Below Beech Trees

To identify the season- and temperature-specific subcanopy micrometeorological effects of preexisting beech dominance within temperate deciduous forest stands, I used another Biological Forecasting and Hindcasting Tool developed by Professor Kearney's group, the Global Soil Microclimate Calculator (Kearney & Porter[5]). For calculations of microclimate conditions (air temperature, humidity, wind speed, and solar radiation), this tool relies on the microclimate model built into NicheMapR, which is itself informed about surface climate by the 1960-1990 global climate grids produced by New et al.[5] The NicheMapR microclimate model consists of a Fortran library, the main calculation engine, and an assisting set of R functions that establish data inputs and call the Fortran program. For further information on the Fortran library, user inputs, and R outputs, see Supplementary Figure 6.

To simulate the impact of beech tree presence on otherwise bare soil in the Global Soil Microclimate Calculator, I manipulated the following terrain and soil parameters which would be most affected by the density of conspecifics in a beech stand: (1) percent shade (2) wind multiplier, and (2) percent albedo. I generated two sets of micrometeorological predictions based on two treatment regimes of these parameters corresponding

to trees-present and bare soil terrain conditions based on the projected impacts of beech forestation throughout the year[7,8,13]. See Supplementary Figure 2 for an example of parameter input and Supplementary Figure 3 for an example of simulation output. I repeated treatments of tree presence and absence for simulations conducted in January and June, with and without a 2 °C climate offset to incorporate seasonal variation and compare the impact of beech tree subcanopy micrometeorology effects with and without macroclimatic warming. A location of (42° N, −79° E), roughly within Allegheny National Forest in the northeastern United States, was used across treatments.

Results

Part 1: Microhabitat Regulation by Blue Mussels

In the maturing juvenile blue mussels subject to a 365-day growth simulation, a distinct delay of puberty and eventual total inhibition of full maturation can be observed with increasing sustained body temperature treatments (Figure 1). Along this same axis of change, a reduction in final wet mass of 99.05% can be noted between 18 °C and 30 °C treatments (Figure 1). Likewise, the proportionally small reproductive buffer (top allocation layer present after puberty) ultimately vanishes with increasing body temperature, as puberty is delayed from day ∼110 to ∼270 until it is fully inhibited by $Tb = 28$ °C (Figure 1).

Part 2: Micrometeorology Below Beech Trees

The effect of beech tree presence on surface micrometeorology depends on season and climate offset. In the January treatment without a climate offset, the presence of trees did not affect minimum soil temperature, minimum air temperatures, minimum humidity, or maximum solar radiation (Table 1). However, with 2 °C of warming, tree presence decreased minimum soil temperatures from 0 °C to −3 °C (Table 1). Regardless of the warming offset, winter tree presence decreased maximum wind speed by 25% (Table 1). In the June simulation, tree presence was not associated with any changes in maximum solar radiation (Table 1). It was associated with a 24-25% reduction in maximum air temperature, depending on the warming scenario 1. Regardless of warming offset, the presence of trees in June increased minimum humidity from 30% to 45%, reduced maximum wind speed by 38%, and reduced maximum soil temperature by 41% (Table 1).

Discussion

Part 1: Microhabitat Regulation by Blue Mussels

The dramatic reduction in total growth, delay of puberty, and loss of the reproduction buffer observed in here in blue mussels with simulated heating from 18 °C to 30 °C reflects their posited ideal body temperature of 16 °C (ref. 17) and their upper limit of temperature tolerance, 23 °C (refs. 16, 17, 20). That being said, the loss of fitness observed here in simulated blue mussels subject to higher temperatures does not necessarily indicate that temperature amelioration will always lead to improved population-level fitness. This is due to the possibility of other ecological constraints (e.g., predation, competition, etc.) not considered in this single species DEB-based niche model potentially constraining the niche of this species more than thermal stress.

However, if this simulation of sustained body temperatures reflects the potential difference in growth and reproductive outcomes between mussels that mature on developed conspecific beds and those that settle on alternative substrates, then it evidences the theoretical positive niche construction ability of blue mussels, wherein the habitat modification performed by small invasive populations may facilitate the niche requirements of future individuals in otherwise unviable habitat[22]. If so, the hypothetical distribution of blue mussels becomes complicated by the existence of an Allee niche. This refers to the region of niche space in which there is a positive association between individual blue mussel fitness and population size[22]. Over time, if populations of blue mussels that settle on cooler developed conspecific beds continue to benefit from this Allee effect, then their invasion of previously unviable habitat may be accelerated following the survival of potentially few pioneering individuals that ameliorate conditions for following conspecifics. As studies have found contemporary ecophysiological models for blue mussels underestimate the spatial extent of their growth when compared to field observations[16], the current framework for mechanistic niche modelling of imposing physiological constraints on target species may systematically underestimate the true extent of habitat potentially suitable to these organisms by ignoring the facilitative microhabitat refugia identified here.

However, to validate the potential existence and longevity of an Allee niche for blue mussels, future research should compare the magnitude of positive, temperature ameliorating, and negative, density-dependent effects (e.g. intraspecific competition, decreased water flow) experienced by mussels living alongside conspecifics. If growth and reproductive outcomes remain higher for mussels on conspecific beds than for lone individuals, then further support would be found for net positive intraspecific interactions and positive niche construction. Conversely, if resource competition and overcrowding within mussel beds are found to limit mussel growth and fecundity in beds still subject to temperature amelioration across various carrying capacities, then positive niche construction, supported in this study, may prove insignificant to the persistence of mussels in the long term due to dominating negative intraspecific interactions.

Part 2: Micrometeorology Below Beech Trees

The summertime effect of beech tree presence on the simulation site can be summarized as an amelioration of hot and dry conditions below the canopy, wherein minimum humidity increased and the maximum temperature of both air and soil decreased. This phenomenon reflects the documented ability of beech stands and forests to generally increase local water availability by increasing the residence time of precipitation stored in biomass and continuously evapotranspiring[1]. Similarly, the significant additional shade provided by a largely closed beech canopy (∼85%) reduces the amount of solar radiation reaching the subcanopy and surface, explaining the reduced maximum temperatures observed here.

The micrometeorological effects of tree presence in January reflect the loss of structural complexity caused by seasonal defoliation. With near-total canopy opening and thus minimal impedance of solar radiation, subcanopy air temperature was not impacted. Likewise, the domination of deciduous trees in this forest would lead to dramatically reduced evapotranspiration during this season, explaining the absence of humidity gains with trees present. The slightly reduced minimum soil temperature observed in forested sites provides an interesting insight into the nontrivial response of soil temperature to the presence of standing stock. Under high snow conditions (0 °C offset), no temperature below 0 °C is recorded. This is likely due to the insulating effect of this snow layer. However, with 2 °C warming, the reduced snow layer may partially vanish near pockets of warmed vegetation, allowing the exposed soil below to be cooled below freezing later on by convection. Thus, the influence of standing stock on soil temperatures may depend on the abundance of other insulating cover, including snow.

The reduced maximum wind speed observed at forested sites in both June and January may be explained by the documented ability of beech stands to increase friction in the boundary layer and dissipate wind below the

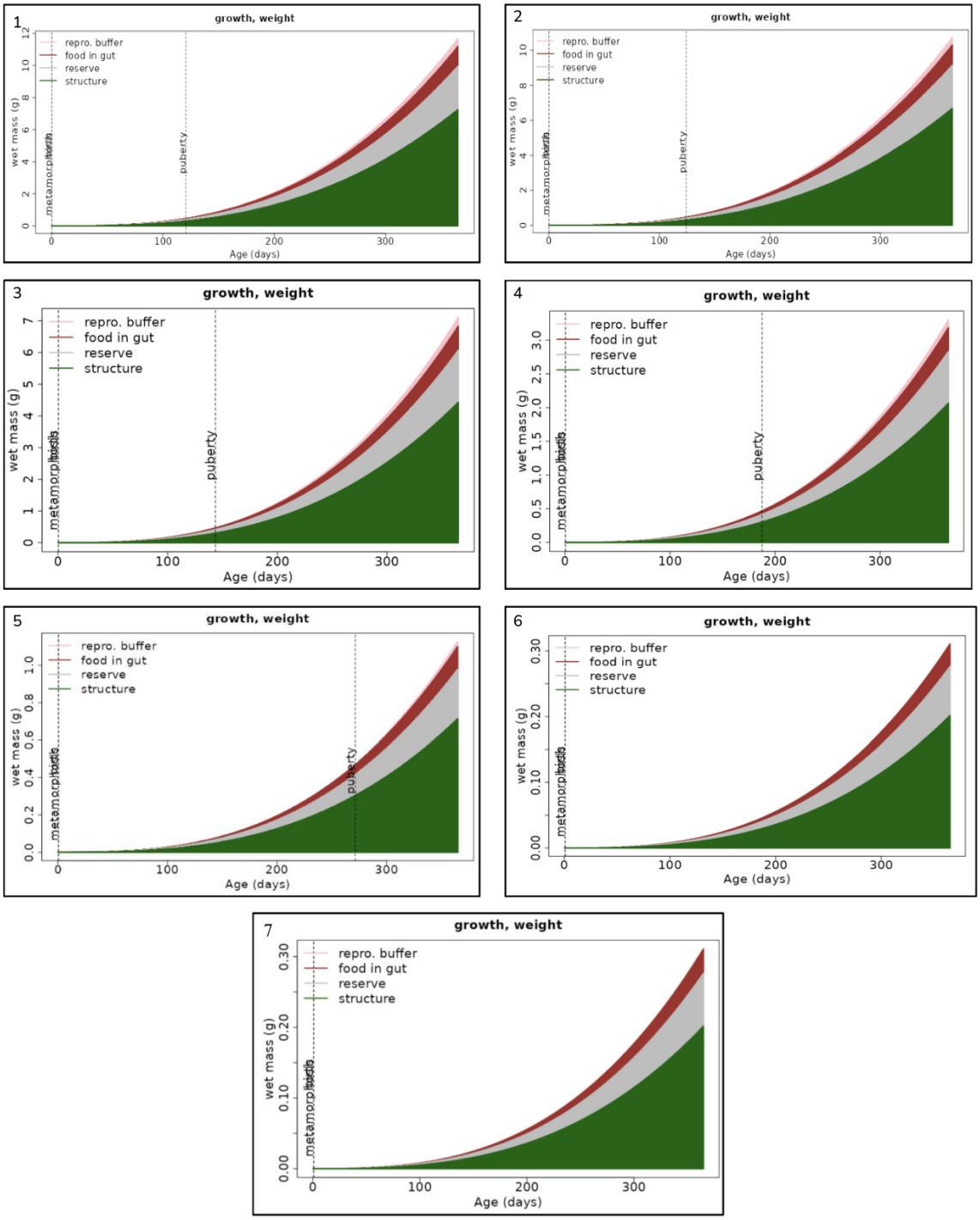


Figure 1. Panels 1-7 correspond to the simulated growth and biomass allocation outputs of DEB in the Sea (Kearney & Porter[5]) simulations for maturing blue mussels at constant body temperatures (1: $Tb = 18\,^{\circ}$C, 2: $Tb = 20\,^{\circ}$C, 3: $Tb = 22\,^{\circ}$C, 4: $Tb = 22\,^{\circ}$C, 5: $Tb = 26\,^{\circ}$C, 6: $Tb = 28\,^{\circ}$C, 7: $Tb = 30\,^{\circ}$C). Note the first dashed line corresponds to birth and immediate metamorphosis, while the second marks the onset of puberty. Parameters of simulation: Location = (54,3), Days = 365, Time step = hourly, Start date = Jan 31, 2020 at 1PM, $z = 4.2$, and $\kappa = 0.95$.

canopy[7]. However, when beech trees evapotranspire, they contribute to a super-subcanopy air density gradient which can create a vertical wind profile[7]. Therefore, the consistency of the forestation effect on maximum wind speed across seasons in these simulations may demonstrate that the physical structure of beech-dominated stands more significantly affects surface layer winds than their seasonally fluctuating evapotranspiring behavior does. The lack of changes in maximum solar radiation across all treatments may be explained by the assumedly incomplete closure of the beech canopy, which is characteristic of non-climax communities and results in gaps through which maximum solar radiation may reach lower vegetation levels. The overall greater amelioration of physical stressors by tree presence in the summer months agrees with the maximal strength of forest influence being recorded during this season[14].

The patterns in micrometeorology that emerge from these simulations of beech stand presence or absence can be used to analyze the comparative impact of cooling or warming and reforestation on surface conditions. Assuming the local terrain and soil property regimes compared in these

Table 1. Table of micrometeorological variables. Derived from the Global Soil Microclimate Calculator (Kearney & Porter[5]), these variables describe sites exhibiting terrain and soil parameters characteristic of bare surface or beech trees present. Outputs reflect daily average conditions below the canopy simulated in June (orange) or January (blue) with and without a +2 °C warming offset. Highlighted green values in Forestation Effect columns reflect notable variation in micrometeorology after the addition of trees to bare sites.

Micrometeorology	Bare Soil	Trees Present	Forestation Effect	Forestation effect (% difference)	Bare Soil +2 °C	Trees Present + 2 °C	Forestation Effect + 2 °C	Forestation effect (% difference)	Month
Max Ts (°C)	44	26	-18	-41%	46	27	-19	-41%	June
Max Ta (°C)	32	24	-8	-25%	34	26	-8	-24%	
Max Wind Speed (m/s)	0.8	0.5	-0.3	-38%	0.8	0.5	-0.3	-38%	
Max Solar Rad (W/m^2)	600	600	0	0	600	600	0	0	
Min Humidity (%)	30	45	15	50%	30	45	15	50%	
Min Ts (°C)	0	0	0	0	0	-3	-3	undefined	January
Min Ta (°C)	-10	-10	0	0	-8	-8	0	0	
Max Wind Speed (m/s)	1	0.75	-0.25	-25%	1	0.75	-0.25	-25%	
Max Solar Rad (W/m^2)	200	200	0	0	200	200	0	0	
Min Humidity (%)	50	50	0	0	50	50	0	0	

Global Soil Microclimate Calculator simulations accurately reflect the effect of beech tree presence on percentage of shade, wind, and albedo in this study area, these results correspond to reasonable predictions of the micrometeorological impact of tree density in beech-dominated forests.

While macroclimate cooling can ameliorate hot and dry conditions, the impact of canopy cover on the subcanopy's micrometeorology has been shown to far outweigh the cooling effect on bare soil. Even with peak canopy openness, the influence of standing stock on microclimate remains observable. Thus, although microclimate researchers may need to incorporate other Earth systems like hydrology into their estimates of local environmental conditions, this simulation demonstrates the applicability of canopy and dominant vegetation data in contemporary DEB-based mechanistic niche models of forest-resident species. To validate the positive niche-construction ability of beech trees, as supported here by the observable amelioration of micrometeorology, further research would need to compare the ecophysiological benefit of a more amenable micrometeorology, as quantified here, with the cost of negative intraspecific effects. In beech trees, these negative effects may include competition for light, water, and nutrients as well as population density-dependent disease vectors. If establishment and recruitment outcomes remain higher for beech trees grown around conspecifics, then further population-level support would be found for net positive niche construction.

Acknowledgements

I thank all members of the Guichard Lab for welcoming me into a collaborative working environment. In particular, I thank Professor Fred Guichard, my supervisor, for codesigning this project's objectives, and PhD student Jake Harvey for his suggestion of contextualizing the results of this study within the niche framework of ecology.

Supplementary Material

Supplementary material referenced in the text of this article may be found online at https://doi.org/10.26443/msurj.v1i1.216.

References

1. Jurgens, L. J., Ashlock, L. W. & Gaylord, B. Facilitation alters climate change risk on rocky shores. *Ecol.* **103,** e03596 (2022). https://doi.org/10.1002/ecy.3596
2. Kenchington, E. et al. *Identification of Species and Habitats that Support Commercial, Recreational or Aboriginal Fisheries in Canada.* Canadian Science Advisory Secretariat (2013).
3. Morrison, M., Consalvey, M., Berkenbusch, K. & Jones, E. Biogenic habitats and their value to New Zealand fisheries. *Water Atmos.* **16,** 20–21 (2008).
4. Lathlean, J. A. et al. Cheating the locals: invasive mussels steal and benefit from the cooling effect of indigenous mussels. *PLoS One* **11,** e0152556 (2016). https://doi.org/10.1371/journal.pone.0152556
5. Kearney, M. R. & Porter, W. P. NicheMapR – an R package for biophysical modelling: the ectotherm and Dynamic Energy Budget models. *Ecography* **43,** 85–96 (2020). https://doi.org/10.1111/ecog.04680
6. Seuront, L., Nicastro, K. R., Zardi, G. I. & Goberville, E. Decreased thermal tolerance under recurrent heat stress conditions explains summer mass mortality of the blue mussel Mytilus edulis. *Sci. Rep.* **9,** 17498 (2019). https://doi.org/10.1038/s41598-019-53580-w
7. Aussenac, G. Interactions between forest stands and microclimate: Ecophysiological aspects and consequences for silviculture. *Ann. For. Sci.* **57,** 287–301 (2000). https://doi.org/10.1051/forest:2000119
8. Holst, T., Mayer, H. & Schindler, D. Microclimate within beech stands? Part II: Thermal conditions. *Eur. J. For. Res.* **123,** 13–28 (2004). https://doi.org/10.1007/s10342-004-0019-5
9. Angelini, A., Corona, P., Chianucci, F. & Portoghesi, L. Structural attributes of stand overstory and light under the canopy. *Ann. Silvic. Res.* **39,** 23–31 (2015). https://doi.org/10.12899/ASR-993
10. Latimer, C. E. & Zuckerberg, B. Forest fragmentation alters winter microclimates and microrefugia in human-modified landscapes. *Ecography* **40,** 158–170 (2017). https://doi.org/10.1111/ecog.02551
11. Kovács, B., Tinya, F. & Ódor, P. Stand structural drivers of microclimate in mature temperate mixed forests. *Agric. For. Meteorol.* **234–235,** 11–21 (2017). http://dx.doi.org/10.1016/j.agrformet.2016.11.268
12. Bequet, R. et al. Leaf area index development in temperate oak and beech forests is driven by stand characteristics and weather conditions. *Trees* **25,** 935–946 (2011). https://doi.org/10.1007/s00468-011-0568-4

13. Frey, S. J. K. et al. Spatial models reveal the microclimatic buffering capacity of old-growth forests. *Sci. Adv.* **2** (2016). https://doi.org/10.1126/sciadv.1501392

14. Baker, T. P., Jordan, G. J. & Baker, S. C. Microclimatic edge effects in a recently harvested forest: Do remnant forest patches create the same impact as large forest areas? *For. Ecol. Manag.* **365,** 128–136 (2016). http://dx.doi.org/10.1016/j.foreco.2016.01.022

15. De Frenne, P. et al. Microclimate moderates plant responses to macroclimate warming. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 18561–18565 (2013). https://doi.org/10.1073/pnas.1311190110

16. Wijsman, J. *Dynamic Energy Budget (DEB) modelBlue mussels (Mytilus edulis): technical report INNOPRO project* English (Wageningen Marine Research, 2019). https://doi.org/10.18174/470139.

17. Rosland, R., Strand, Ø., Alunno-Bruscia, M., Bacher, C. & Strohmeier, T. Applying Dynamic Energy Budget (DEB) theory to simulate growth and bio-energetics of blue mussels under low seston conditions. *J. Sea Res.* **62,** 49–61 (2009). http://dx.doi.org/10.1016/j.seares.2009.02.007

18. Monaco, C. J. & McQuaid, C. D. Applicability of Dynamic Energy Budget (DEB) models across steep environmental gradients. *Sci. Rep.* **8,** 16384 (2018). https://doi.org/10.1038/s41598-018-34786-w

19. Saraiva, S., Der Meer, J. V., Kooijman, S. A. L. M. & Sousa, T. DEB parameters estimation for Mytilus edulis. *J. Sea Res.* **66,** 289–296 (2011). https://doi.org/10.1016/j.seares.2011.06.002

20. Saraiva, S. et al. Validation of a Dynamic Energy Budget (DEB) model for the blue mussel Mytilus edulis. *Mar. Ecol. Prog. Ser.* **463,** 141–158 (2012). https://doi.org/10.3354/meps09801

21. Van Der Veer, H. W., Cardoso, J. F. M. F. & Van Der Meer, J. The estimation of DEB parameters for various Northeast Atlantic bivalve species. *J. Sea Res.* **56,** 107–124 (2006). https://doi.org/10.1016/j.seares.2006.03.005

22. Koffel, T., Daufresne, T. & Klausmeier, C. A. From competition to facilitation and mutualism: a general theory of the niche. *Ecol. Monogr.* **91,** e01458 (2021). https://doi.org/10.1002/ecm.1458

Submitted: 01/14/2024
Accepted: 03/14/2025
Published: 03/31/2025

Article Type

[1]Department of Cell Biology and Anatomy, McGill University, Montréal, QC Canada

[2]Research Institute of the McGill University Health Centre, Montréal, QC, Canada

[3]Department of Human Genetics, McGill University, Montréal, QC, Canada

[4]Division of Medical Genetics, McGill University Health Centre, Montréal, QC, Canada

Keywords

Primary ciliary dyskinesia, Heart laterality, Morpholino-oligonucleotide injections, Zebrafish, Genetics

Email Correspondence

divya.kakkar@mail.mcgill.ca

https://doi.org/10.26443/msurj.v20i1.222

Divya Kakkar[1], Zachary W. Nurcombe[2,3], Lina Mougharbel[2] and Thomas M. Kitzler[2, 3, 4]

Assessment of Heart Laterality Defects in Zebrafish to Study Variants of Uncertain Significance in Primary Ciliary Dyskinesia

Abstract

Primary ciliary dyskinesia (PCD) is an autosomal recessive orphan disease (OMIM#244400) characterized by motile ciliary dysfunction. These hairlike organelles are responsible for the mucociliary clearance of the lungs, and varying degrees of infections in the upper respiratory tract—including the inner ear, nasal passage, and lungs—are common in affected children present with PCD. Moreover, ciliary function is critical for embryonic development, and defects in cilia can lead to situs anomalies, which are sometimes associated with congenital heart disease. To date, more than 50 genes have been implicated in the etiology of PCD, each affecting different parts of the motile ciliary apparatus. Testing via multi- or single-gene panels is recommended for confirmation of diagnosis, which enables timely treatment initiation and familial risk counselling. Unfortunately, for a significant proportion of children with clinical features consistent with PCD, a molecular diagnosis cannot be established. For many of these children, genetic testing returns a variant of uncertain significance (VUS) in a known PCD gene. Hence, there is a pressing need to develop strategies to validate unresolved PCD variants for their pathogenicity. The overall aim of this project is to resolve VUSs in children with suspected but genetically unconfirmed PCD by use of zebrafish (*Danio rerio*). As a key PCD gene in humans, *dnaaf1* has a *dnaaf1* homolog in zebrafish, making this model highly relevant for studying human PCD phenotypes. By employing gene knock-down technology (antisense morpholino oligonucleotides; MO), we established and characterized PCD-specific developmental readouts for the zebrafish gene *dnaaf1*, which affect different aspects of motile ciliary ultrastructure. We observed an increase in ventral body curvature and hydrocephalus in embryos with *dnaaf1*-MO, with 38% of *dnaaf1*-MO knockdown embryos showing heart-laterality defects. These phenotypic outcomes not only provide a concrete framework for assessing PCD-related developmental defects in zebrafish but also offer a platform for validating VUSs in human PCD genes. By performing co-injection experiments with patient-derived VUSs and examining the resulting phenotypic alterations, we can directly link specific genetic variants to observable PCD-like traits, offering a robust methodology for determining the pathogenicity of previously uncharacterized variants. This approach aims to enhance the accuracy of genetic diagnoses in PCD and provide new insights into its molecular mechanisms.

Introduction

Motile ciliary dysfunction defines the orphan disease (OMIM#244400) primary ciliary dyskinesia (PCD), which follows an autosomal recessive inheritance pattern. PCD is a rare, heterogeneous disease associated with more than 50 monogenic causes resulting from a loss of function in the motile cilia apparatus[1]. Cilia are complex structures encoded by many genes in various cell types and play a significant physiological role in humans. Motile cilia are responsible for the mucociliary clearance in the lungs and is vital in embryonic development. During development, cilia directionally beat to guide the flow of embryonic fluid, which establishes proper left/right axis body asymmetry and organ development. Structural or functional abnormalities in the cilia are formed by variants in known PCD genes. This can lead to two forms of *situs inversus*, conditions in which the normal asymmetry of internal organs is reversed or disrupted. The first, *situs inversus totalis*, involves major organs, including the heart and liver, being mirrored to the opposite side of the body. The second, *situs inversus totalis*, is characterized by irregular or inconsistent organ positioning, resulting in complex and often unpredictable arrangements[2]. These ciliary defects not only disrupt the typical left-right body asymmetry but also impair mucociliary clearance, which can lead to chronic respiratory issues and abnormal embryonic development. As a result, ciliary defect individuals often present with recurrent respiratory infections, such as bronchitis and congenital heart disease. More specifically, about 50% of patients with *situs inversus*—whether *totalis* or *ambiguous*—are also affected by these associated complications, including bronchitis and heart defects[3] (Figure 1). Understanding variants in PCD genes can help elucidate the genetic basis of these conditions and potentially inform diagnostic and therapeutic strategies.

PCD is a childhood onset disease with an estimated prevalence of 1:7,500 to 1:20,000 in liveborn babies[1]. However, this may be an underestimate due to variable clinical presentation and a current absence of readily available diagnostic markers for PCD. A diagnosis of PCD can be obtained clinically by demonstration of structurally abnormal cilia via transmission electron microscopy (TEM) and/or low nitric oxide (NO) inhalation measurements[4]. However, TEM is costly, and not all patients exhibit observable cilia abnormalities. Additionally, nasal nitric oxide (NO) tests are limited to children above the age of 5 years[5]. Therefore, molecular testing via multi-gene panels is the recommended clinical mainstay for early and accurate confirmation of a PCD diagnosis. Identification of the underlying molecular defect enables early supportive treatment initiation, accurate familial and reproductive risk counseling, as well as access to novel personalized treatments. Unfortunately, for a significant number of children, the molecular diagno-

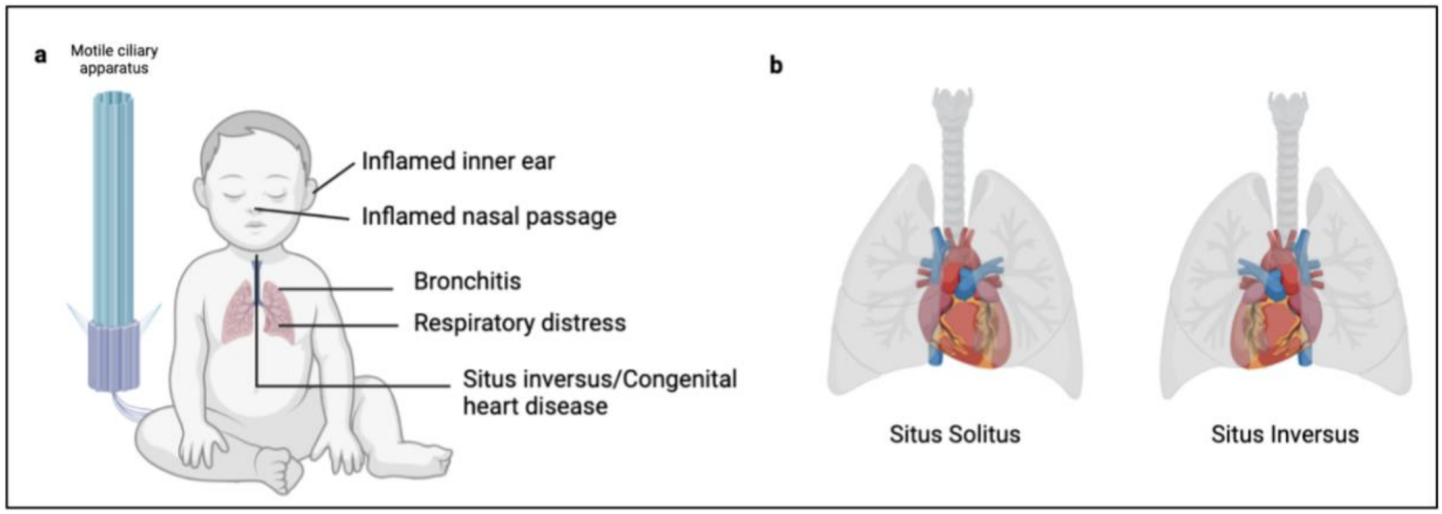


Figure 1. Clinical manifestations of primary ciliary dyskinesia (PCD) and its impact on organ laterality. **(a)** Common symptoms of PCD, including inflammation of the inner ear and nasal passages, bronchitis, respiratory distress, and *situs inversus* with or without congenital heart disease, are shown alongside a schematic of a motile ciliary apparatus. **(b)** Comparison of normal organ positioning (*situs solitus*) versus mirror-image organ arrangement (*situs inversus*), a hallmark phenotype observed in approximately 50% of PCD patients. Created in BioRender. Nurcombe, Z. (2025) https://BioRender.com/1czt0of

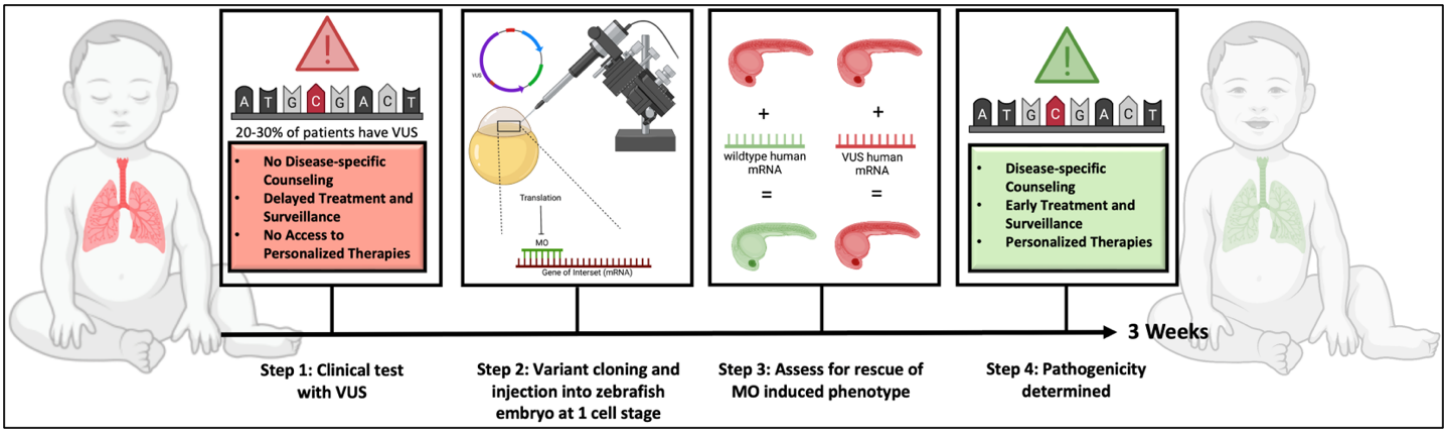


Figure 2. Workflow for assessing the pathogenicity of variants of uncertain significance (VUS) in primary ciliary dyskinesia (PCD) using a zebrafish model. Step 1: Clinical testing identifies patients with VUS, leading to challenges in treatment and counseling. Step 2: The VUS is cloned and injected into zebrafish embryos at the one-cell stage. Step 3: Phenotypic rescue is assessed by co-injecting wild-type and VUS mRNA. Step 4: Pathogenicity is determined, enabling disease-specific counseling, early treatment, and access to personalized therapies within approximately three weeks. Created in BioRender. Nurcombe, Z. (2025) https://BioRender.com/19xd4am

sis remains inconclusive due to a variant of uncertain significance (VUS) in PCD genes. Variants of uncertain significance are point mutations or deletions with unknown impact on gene and protein function, leading to undetermined pathogenicity. We propose zebrafish as a tool to rapidly validate VUS in PCD genes (Figure 2).

This project focuses on the study of the human ciliary PCD gene *Dynein Axonemal Assembly Factor 1* (*dnaaf1*). *dnaaf1* encodes a protein involved in the preassembly and transport of dynein arms, which are essential for the proper function of motile cilia. Dynein arms are motor protein complexes responsible for the coordinated beating of cilia, enabling functions such as mucociliary clearance in the respiratory tract and left-right body asymmetry establishment during embryonic development. Mutations in *dnaaf1* are associated with PCD, leading to defects in dynein arm assembly and loss of ciliary motility. Clinically, variants in *dnaaf1* lead to problems affecting the auditory (hearing loss), cardiovascular (dextrocardia), respiratory (sinusitis, bronchitis, bronchiectasis), and genitourinary (infertility in both male and female at birth) systems in patients. Most notably, 50% of patients with variants in *dnaaf1* present with *situs inversus* where the normal left-right asymmetry of internal organs is reversed[6,7].

The goal of this project is to confirm suspected PCD in patients with an inconclusive molecular diagnosis due to VUS in PCD-related genes by using zebrafish (*Danio rerio*) models. To achieve this, we developed an antisense mRNA probe for *cmlc2*, a gene essential for zebrafish heart development and a well-established marker for cardiac laterality. Since *dnaaf1* is crucial for dynein arm assembly in motile cilia, its disruption can impair ciliary function, leading to defects in left-right patterning during embryonic development. By using the *cmlc2* probe to visualize heart position in zebrafish embryos, we can assess laterality defects, which serve as an indicator of disrupted ciliary function. This approach allows us to determine whether specific *dnaaf1* variants impact motile cilia-driven processes, thereby linking genetic variants to observable phenotypes. Ultimately, this strategy provides functional validation of VUS in *dnaaf1* and other PCD genes, improving diagnostic accuracy for patients with unresolved genetic findings.

Zebrafish serve as an excellent model for establishing PCD phenotypes due to their high genetic homology (~70%) with humans, including the conservation of key genes such as *dnaaf1*, which is critical for ciliary function and motile cilia-driven processes[8]. One of the major advantages of zebrafish as a model organism is their rapid development and transparency during early

larval stages, allowing for real-time visualization of internal structures under light microscopy. Importantly, zebrafish develop a beating heart as early as 24 hours post-fertilization (hpf), providing a fast and accessible readout for assessing laterality defects associated with impaired ciliary function. Since ciliary dysfunction in PCD disrupts left-right asymmetry during embryonic development, the early formation of the zebrafish heart provides a rapid model for assessing whether defective cilia affect organ positioning—a hallmark phenotype of PCD. This rapid timeline enables efficient screening of genetic variants and accelerates functional validation of VUS in PCD-related genes. Overall, observing heart laterality in zebrafish is particularly valuable for investigating developmental pathways associated with PCD and translating these insights to human disease[9].

The *cmlc2* gene was used to determine heart laterality defects in zebrafish by serving as a molecular marker for cardiac tissue, enabling precise visualization of heart positioning. *cmlc2* is specifically expressed in the developing heart, making it an ideal target for assessing left-right patterning abnormalities associated with PCD-related ciliary dysfunction. To analyze heart laterality, we designed an antisense mRNA probe targeting *cmlc2* and used whole mount in situ hybridization (WISH) to stain and visualize the zebrafish heart. This allowed us to determine whether the heart was positioned normally (*situs solitus*, SS), abnormally mirrored (*situs inversus*, SI), located at the midline (midline heart, M), or duplicated (bilateral heart, B), all of which are key phenotypes linked to defective cilia function.

Morpholino knockdown technology is a widely used tool in zebrafish research to study gene function during development. Morpholinos are synthetic antisense oligonucleotides designed to bind complementary sequences in target mRNA, thereby blocking either translation or proper splicing. Unlike traditional RNA interference (RNAi), which relies on cellular machinery to degrade mRNA, morpholinos work by blocking ribosome access, preventing protein synthesis without triggering mRNA degradation. This makes them particularly useful for gene knockdown in early developmental stages.

By injecting morpholinos into zebrafish embryos, we were able to knock down known PCD-related genes and establish their corresponding phenotypes. This approach has already been successfully used to study *dnaaf1*, *dnaaf3*, and *dnah1* (Figure 1), revealing hallmark PCD phenotypes such as perturbed otoliths, hydrocephalus, body curvature defects, ciliary abnormalities in the eye, pronephric cysts, and *situs inversus*[9]. Notably, heart laterality defects—such as SS, SI, M, and B—have been well documented in *dnaaf1* and *dnaaf3* knockdown models[10,11]. Additionally, body curvature defects have been associated with *dnaaf1* loss, further confirming their role in cilia-driven developmental processes. Both heart laterality and body curvature abnormalities stem from defective ciliary function. Given the significance of these phenotypes associated with *dnaaf1* loss, we aimed to evaluate the pathogenicity of variants in *dnaaf1* through body curvature and hydrocephaly phenotypes. These two phenotypes are characteristic of PCD in zebrafish and therefore are a confident readout for these experiments.

To assess heart laterality defects in zebrafish models of PCD, researchers commonly probe for *cmlc2* (cardiac myosin light chain 2), a gene specifically expressed in the developing heart. *cmlc2* serves as a molecular marker for cardiac tissue, allowing visualization of heart positioning relative to the embryonic midline. By examining *cmlc2* expression patterns, we can determine whether the heart is properly positioned (SS) or exhibits abnormal laterality, such as leftward displacement (SI) or ambiguous positioning, all key phenotypes indicative of ciliary dysfunction.

Objective

We developed a probe to enable characterization of a PCD-specific phenotype in zebrafish. And we assessed heart laterality defects associated with mutations in the PCD gene, *dnaaf1*, to help confirm suspected PCD in patients with VUS.

Materials and Methods

Microinjections

Morpholinos (200 ng) (*Ensembl*) are injected to knock down gene expression in zebrafish and create knockdowns of targeted genes for up to 5 days post-fertilization (dpf)[12,13]. The *dnaaf1*-MO (Table 1a) inhibits expression of *dnaaf1* in zebrafish through a translation-blocking mechanism. Zebrafish embryos were injected (5 ng) at the 1-2 cell stage with 28.5 ˚C incubation, with either a non-targeting MO (Table 1b) as a negative control (250 uM) or a *dnaaf1*-MO (250 µmM), generating *dnaaf1* morphant embryos to establish the PCD-specific phenotypes in zebrafish. Embryos were fixed at 24 hpf in 4% paraformaldehyde (PFA).

a	*dnaaf1*-MO	5'-ATTTTTGTCTTCATTTCGCAGTGAT-3'
b	scrambled MO	5'-CCTCTTACCTCAGTTACAATTTATA-3'
c	PCR primers for *cmlc2* probe synthesis	f: 5' – GACCAACAGCAAAGCAGACA – 3' r: 5' – TAATACGACTCACTATAGGG GGGTCATTAGCAGCCTCTTG – 3'

Table 1. Sequences for PCR primers and MOs. a) *dnaaf1*-MO sequence. b) scrambled control MO sequences. b) PCR primers for *clmc2* probe synthesis. T7 RNA polymerase binding site in red.

Probe Synthesis

Designing an antisense mRNA probe for the *cmlc2* gene allows visualization of the developing zebrafish heart to assess heart laterality, a component of the PCD phenotype. The *cmlc2* gene contains a 77% similarity to the human homolog *MYL2*[14,15]. This zebrafish ortholog is a small transcript of roughly 1000 bp (NCBI Reference Sequence: NM_131329.3) and is expressed in the developing zebrafish heart at 24 hpf. To target a region of the *cmcl2* transcript (under 500 bp), forward and reverse primers for PCR were designed against the zebrafish *cmlc2* coding sequence. The reverse primer includes the TAATACGACTCACTATAGGG sequence, which serves as the T7 RNA polymerase binding site (Table 1c). PCR was carried out using 48 hpf zebrafish cDNA as the template, with Platinum SuperFi II polymerase according to manufacturer guidelines. This process ensured the incorporation of the T7 site at the 5' end of the *cmlc2* amplicon, enabling *in vitro* transcription. This T7 sequence allowed T7 RNA polymerase to synthesize antisense mRNA, which was used as the probe for detecting *cmlc2* expression. The alkaline phosphatase conjugated anti-digoxigenin (α-DIG) antibody reacts with the NBT/BCIP, which produces the purple signal of the probe in the zebrafish heart. The NBT/BCIP solution consisted of 4.5 mM nitro-blue tetrazolium chloride (NBT), 3.75 mM 5-bromo-4-chloro-3-indolyl phosphate (BCIP), 100 mM Tris buffer (pH 9.5), 50 mM magnesium chloride ($MgCl_2$), and 0.1% Tween-20. The DIG-labeled antisense mRNA probe hybridized to its complementary target in the zebrafish heart and was detected by an alkaline phosphatase (AP)-conjugated anti-digoxigenin (α-DIG) antibody. The colorimetric reaction occurs when NBT/BCIP, added in the staining buffer, serves as a substrate for AP, leading to the formation of an insoluble purple precipitate at the probe's location. This occurs because the antisense mRNA is labeled with digoxigenin (DIG)-UTPs during

in vitro transcription, allowing the α-DIG antibody to specifically bind to the probe and facilitate the colorimetric reaction.

Whole Mount mRNA In Situ Hybridization (WISH)

Standard WISH in zebrafish against the *cmlc2* transcript was carried out to visualize the heart, allowing us to assess laterality defects[16]. We fixed the embryos at 24 hpf in 4% PFA overnight at 4 ˚C. Fixed embryos were then dehydrated in 100% methanol and stored at –20 ˚C. Embryos were rehydrated through a series of methanol-phosphate buffered saline (PBS) washes (75/25, 50/50, 25/75, 0/100). The *cmlc2* probe (1:100 stock probe in hybridization buffer – formamide 50%, 5xSSC, Tween 20 0.1%, heparin 50ug/ml, torula RNA 500ug/ml, pH 6.0) is placed on the embryos and then are incubated at 70 ˚C overnight. Next, the embryos underwent two 30-minute washes in 2xSSC (saline sodium citrate buffer) at 70 ˚C and two 30-minute washes in 0.2xSSC at room temperature (RT). We then blocked the embryos with blocking buffer (PBT containing 2% sheep serum and 2 mg/ml bovine serum albumin (BSA)) for 1 hour at RT, followed by a 2-hour incubation with the anti-digoxigenin (α-DIG) antibody (diluted 1:1000 in blocking solution). After antibody removal, we incubated the embryos in PBT (PBS with 1% Tween 20) overnight at 4 ˚C. The embryos were washed twice in staining buffer (0.1 M Tris HCl pH 9.5, 0.1 M NaCl) for 15 minutes to eliminated residual PBS/PBT. Subsequently, they are diluted by adding NBT/BCIP (nitro blue tetrazolium/5-bromo-4-chloro-3-indolyl-phosphate) stock solution to the staining buffer (100 µm in 5 ml) and then added to embryos. Embryos were kept in the dark for roughly 30 minutes allowing time for the signal to develop. Once the signal had developed, we post-fixed the embryos in 4% PFA for 20 minutes at RT. The embryos then underwent a final wash and were stored in PBS at 4 ˚C for subsequent processing. We followed the Thisse and Thisse protocol[16].

Imaging

For assessment of body curvature and hydrocephalus, we live-imaged the embryos laterally at 48 hpf using basic light microscopy (LeicaMZ7.5 microscope with Opti Vision 4K Series CMOS microscope camera). For heart laterality, the embryos were fixed at 24 hpf, stained, and then individually mounted in 3% methylcellulose and imaged cephalically (Figure 2a) using basic light microscopy.

Quantification

To avoid bias during the assessment of heart position, the conditions were blinded during the quantification process. In zebrafish, heart position is classified in terms of its relative location to the eyes. During development, the heart initially forms at the midline and then loops to its designated side by approximately 48 hpf (ref. 17) (Figure 3a). The heart positions are classified as follows: If the heart is located from the lateral wall to the medial edge of the left eye, it is classified as *situs solitus* (SS). If the heart is located under the lateral wall to the medial edge underneath the right eye, it is classified as *situs inversus* (SI). Located between the two medial edges of the eyes is a midline heart (M) and if there are two hearts present it is classified as a bilateral heart (B). The embryos were appropriately quantified under the four conditions (SS, SI, M, B) and then subjected to further data analysis to assess the relevant phenotypic outcomes.

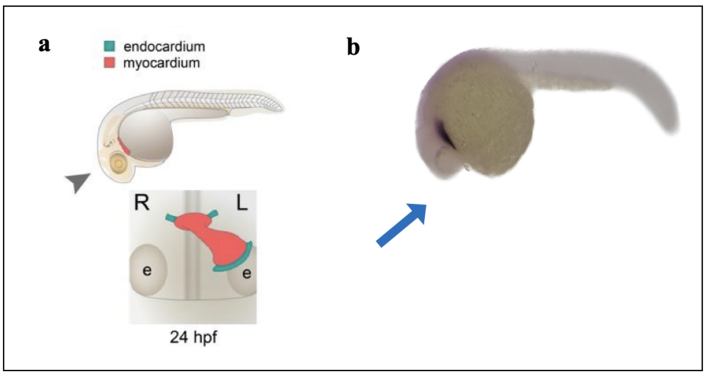


Figure 3. Imaging and quantification of zebrafish embryos at 24 hpf. a) Orientation of embryo for cephalic imaging via light microscopy of embryos fixed at 24 hpf to properly observe heart laterality b) Normal heart development position in zebrafish embryos at 24 hpf. Heart anchored medially with looping located directly under the left eye[17].

Figure 4. Body curvature phenotype in *dnaaf1*-MO knockdown embryos. a) Lateral view of *dnaaf1*-MO knockdown embryos at 48 hpf via live imaging exhibiting increased body curvature and hydrocephalus (red arrow).

Results

Body Curvature

To assess the role of *dnaaf1* in early zebrafish development, we performed morpholino (MO) knockdown and analyzed embryonic phenotypes at 48 hours post-fertilization (hpf). Compared to scrambled MO controls, *dnaaf1*-MO knockdown embryos exhibited pronounced ventral body curvature and increased hydrocephalus, suggesting defects in body axis formation and ciliogenesis. These phenotypic abnormalities are consistent with disrupted motile cilia function, which is required for proper left-right patterning and cerebrospinal fluid circulation. As shown in Figure 4a, the scrambled MO control embryo (top panel) displayed a normal elongated body axis and typical brain morphology. In contrast, the *dnaaf1*-MO knockdown embryo (bottom panel) exhibited a curved body axis and an enlarged brain ventricle indicative of hydrocephalus (red arrow). Quantification revealed that *dnaaf1*-MO embryos (n=17) presented with body curvature defects, supporting the role of *dnaaf1* in maintaining normal body symmetry. Statistical analysis was performed using a Chi-square test ensuring an appropriate comparison of categorical data. Replicates consisted of individual embryos from at least three independent experiments to confirm reproducibility.

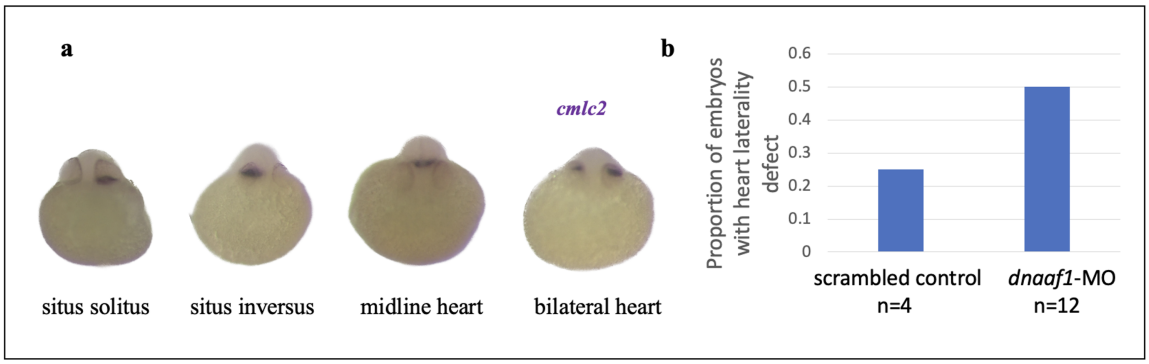


Figure 5. Laterality defect phenotypes and proportions in MO knockdown embryos. a) Cephalically imaged *dnaaf1*-MO knockdown embryos showing heart laterality phenotypes: *situs solitus* (normal), *situs inversus*, midline heart, and bilateral heart. b) Proportion of *dnaaf1*-MO knockdown embryos demonstrating heart laterality defects (0.14 versus 0.38; P=0.027).

Heart Laterality Defects

To evaluate heart laterality defects in *dnaaf1*-MO knockdown embryos, we used WISH with a *cmlc2* antisense probe to visualize *cmlc2* mRNA expression in cardiac tissue. *cmlc2* is a well-established marker of the zebrafish heart, allowing for precise localization and assessment of cardiac positioning in developing embryos. By comparing the heart's positioning in *dnaaf1*-MO embryos versus scrambled MO controls, we were able to identify defects in left-right asymmetry, a hallmark of PCD-related ciliopathy.

As shown in Figure 5a, we observed four distinct heart laterality phenotypes: *situs solitus* (normal), *situs inversus*, midline heart, and bilateral heart. Compared to controls, *dnaaf1*-MO knockdown embryos exhibited a significantly higher incidence of heart laterality defects, with a noticeable increase in *situs inversus*, midline heart, and bilateral heart phenotypes. Quantification of heart laterality defects (Figure 5b) revealed that 38% of *dnaaf1*-MO embryos (n=17/45) displayed abnormal heart positioning compared to only 14% of scrambled MO controls (n=7/50), with a statistically significant difference (P=0.027, Chi-square test). This analysis was conducted using biological replicates from at least three independent experiments, ensuring statistical robustness. Additionally, each n value represents individual embryos, with percentages calculated per replicate rather than as a single data point. These findings support the role of *dnaaf1* in proper left-right patterning.

Discussion

By morpholino mediated gene knockdown, we observed PCD phenotypes in 48 hpf zebrafish embryos. The first PCD phenotype observed is an increase in ventral body curvature in *dnaaf1*-MO knockdowns (Figure 3a). This is consistent with the primary PCD phenotype of ventral body curvature in zebrafish for the PCD gene *Lrcc50*, observed by Van Rooijen et al. (2009)[10]. These authors developed a line of zebrafish with a mutation in the *Lrcc50* gene. Their mutant embryos showed a pronounced ventral body curve, and cilia of Kupffer vesicles, which are required to break bilateral symmetry in zebrafish, showed no motility. Mutant pronephric tubules showed cysts and reduced brush border with short microvilli. Electron microscopy of pronephric mutant cilia revealed ultrastructural irregularities, and some axonemes completely lacked all dynein arms or had misplaced inner dynein arms. Van Rooijen et al. (2008) concluded that *Lrcc50* is critical for ciliary architecture, proper left-right asymmetry, organ positioning, and body axis curvature. Improper cilia function leads to malformation of the curvature of the embryos, making it a good marker to characterize PCD. Their methodology is congruent with ours in using zebrafish KOs as a model of disease. However, their experiments were conducted with 24 hpf embryos, different from our 48 hpf time point. The difference in developmental stages is an important consideration. At 24 hpf, zebrafish are still in the early stages of body plan establishment, and defects in body axis curvature may be more subtle or in an earlier phase of manifestation. In contrast, by 48 hpf, the curvature defects may be more pronounced and fully realized, which could explain the more obvious phenotypic observations in our study.

The other observed PCD phenotype from our study is an increase in heart laterality defects in *dnaaf1*-MO knockdown embryos at 48 hpf. The *dnaaf1*-MO embryos demonstrated *situs solitus* (SS), *situs inversus* (SI), midline heart (M), or bilateral heart (B) (Figure 4a). These heart laterality defects are a hallmark of ciliary dysfunction, and they are commonly associated with PCD in both zebrafish models and humans. The findings of our study align with the work of Van Rooijen et al. (2009)[10], who also reported similar heart laterality defects in their own PCD morphants zebrafish. Van Rooijen et al. (2009)[10] observed all heart laterality phenotypes in *dnaaf1* morphants except the bilateral heart phenotype. The cause of the bilateral heart phenotype is unknown and requires further investigation, however we still considered it as a laterality defect for quantification of this experiment.

In humans, mutations in *dnaaf1*, a gene involved in the assembly of dynein arms in motile cilia, lead to a variety of PCD-related defects, including heart laterality defects. Approximately 50% of individuals with *dnaaf1* mutations exhibit heart laterality defects, including *situs inversus* or situs *ambiguous* (partial *situs inversus*), in which the heart is either reversed or positioned incorrectly due to defects in left-right patterning[18]. As expected, we observed a statistically significant (P=0.027) proportion of heart laterality defects (Figure 4b), confirming its specificity for PCD. This result confirms the specificity of heart laterality defects as a reliable marker for studying PCD and further validates the zebrafish model for investigating the molecular mechanisms of this disease.

While MO knockdown remains a widely used approach for gene silencing in zebrafish, we must acknowledge its potential for off-target effects, which have been extensively documented in the field. To mitigate this, we ensured specificity by using a scrambled MO control and analyzing phenotypic consistency across multiple independent experiments. However, future studies using CRISPR-Cas9 or mutant lines would be valuable in further validating the role of *dnaaf1* in early zebrafish development.

Conclusion

The goal of this study was to develop a heart-specific probe and establish the heart laterality defect as a PCD-specific phenotype in zebrafish, par-

ticularly in the context of *dnaaf1* knockdown. We utilized MO-mediated gene knockdown technology to disrupt *dnaaf1* expression in zebrafish embryos, with the hypothesis that these embryos would exhibit an increased proportion of heart laterality defects, a hallmark of PCD. Through our experiments, we demonstrated that *dnaaf1*-MO knockdown embryos did indeed show an increased incidence of heart laterality defects, including *situs inversus*, midline heart, and bilateral heart phenotypes, providing strong evidence that *dnaaf1* is a key player in the establishment of left-right asymmetry and proper organ positioning in zebrafish. These findings confirm the utility of the heart laterality defect as a reliable and specific phenotype for PCD in zebrafish models.

Our findings contribute to the ongoing project aimed at establishing zebrafish as a robust *in vivo* model for the study of PCD. The overall aim is to establish zebrafish as a tool for rapid variant validation to confirm probable PCD diagnoses. By optimizing MO-knockdown protocols and the heart-specific probe we developed, our research lays the groundwork for future studies focused on the role of other PCD-associated genes, such as *dnaaf3* and *hydin*, where phenotypic characterization in zebrafish has yet to be established. These additional gene knockdowns will help broaden our understanding of the diverse phenotypic spectrum of PCD and contribute to the genetic and clinical characterization of this complex disease.

Acknowledgements

Thank you to the Kitzler lab and Foundation Grand Défi Pierre Lavoie.

References

1. Kuehni, C. et al. Factors influencing age at diagnosis of primary ciliary dyskinesia in European children. *Eur. Respir. J.* **36,** 1248–1258 (2010). https://doi.org/10.1183/09031936.00001010
2. Dasgupta, A. & Amack, J. D. Cilia in vertebrate left–right patterning. *Philos. Trans. R. Soc. B: Biol. Sci* **371,** 20150410 (2016). https://doi.org/10.1098/rstb.2015.0410
3. Escudier, E., Roger, G. & Coste, A. Intérêt du prélèvement nasal pour le diagnostic des dyskinésies ciliaires primitives de l'enfant [Nasal ciliary investigations for the diagnosis of primary ciliary dyskinesia in children]. *Arch. Pédiatr.* **11,** 390–393 (2004). https://doi.org/10.1016/j.arcped.2003.11.030
4. Shapiro, A. J. et al. Diagnosis of primary ciliary dyskinesia. An official American Thoracic Society clinical practice guideline. *Am. J. Respir. Crit. Care Med.* **197,** e24–e39 (2018). https://doi.org/10.1164/rccm.201805-0819ST
5. Leigh, M. W. et al. Standardizing nasal nitric oxide measurement as a test for primary ciliary dyskinesia. *Ann. Am. Thorac. Soc.* **10,** 574–581 (2013). https://doi.org/10.1513/AnnalsATS.201305-110OC
6. *Clinical synopsis table* - #606763 - *OMIM*. OMIM. (n.d.).
7. *Clinical synopsis table* - #613193 - *OMIM*. OMIM. (n.d.).
8. Burke, E. Why use zebrafish to study human diseases? (2016); https://irp.nih.gov/blog/post/2016/08/why-use-zebrafish-to-study-human-diseases
9. Singleman, C. & Holtzman, N. G. Growth and maturation in the zebrafish, Danio rerio: a staging tool for teaching and research. *Zebrafish* **11,** 396–406 (2014). https://doi.org/10.1089/zeb.2014.0976
10. Van Rooijen, E. et al. Zebrafish mutants in the von Hippel-Lindau tumor suppressor display a hypoxic response and recapitulate key aspects of Chuvash polycythemia. *Blood* **113,** 6449–6460 (2009). https://doi.org/10.1182/blood-2008-07-167890
11. Mitchison, H. M. et al. Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia. *Nat. Genet.* **44,** 381–389 (2012). https://doi.org/10.1038/ng.1106
12. Gene Tools. Custom Morpholinos, Controls, and End Modifications (n.d.); https://www.gene-tools.com/custom_morpholinos_controls_endmodifications#standardcontrols
13. Timme-Laragy, A. R., Karchner, S. I. & Hahn, M. E. in *Developmental Toxicology: Methods and Protocols* (eds Harris, C. & Hansen, J. M.) 51–71 (Springer, 2012). https://doi.org/10.1007/978-1-61779-867-2_5
14. U.S. National Library of Medicine. Danio rerio myosin, light chain 7, regulatory (MYL7), mrna – nucleotide – NCBI. *National Center for Biotechnology Information* https://www.ncbi.nlm.nih.gov/nuccore/402693353 (n.d.).
15. InVivo Biosystems. Helio Ortholog Gene Finder (2023); https://invivobiosystems.com/resources/helio-ortholog-gene-finder/
16. Thisse, B. & Thisse, C. Zfin the Zebrafish Information Network (n.d.); https://zfin.org/ZFIN/Methods/ThisseProtocol.html
17. Smith, K. A. & Uribe, V. Getting to the heart of left–right asymmetry: contributions from the zebrafish model. *J. Cardiovasc. Dev. Dis.* **8,** 64 (2021). https://doi.org/10.3390/jcdd8060064
18. Knowles, M. R., Daniels, L. A., Davis, S. D., Zariwala, M. A. & Leigh, M. W. Primary ciliary dyskinesia. Recent advances in diagnostics, genetics, and characterization of clinical disease. *Am. J. Respir. Crit. Care Med.* **188,** 913–922 (2013). https://doi.org/10.1164/rccm.201301-0059CI

Submitted: 01/17/2024
Accepted: 03/06/2025
Published: 03/31/2025

Research Article

[1]Department of Geography, McGill University, Montréal, QC, Canada

[2]Bieler School of Environment, McGill University, Montréal, QC, Canada

Keywords

Feed security, Dairy cattle diets, Québec dairy industry, Dairy land use, Climate adaptation

Email Correspondence

cassia.attard@mail.mcgill.ca

https://doi.org/10.26443/msurj.v20i1.225

Cassia Attard[1,2], Jasmine Boucley[1,2], Elianta Jaillet[1,2], Maxime Léger[1,2] and Eleanor Morrison[1,2]

Assessing Changes in Feed Security of the Québec Dairy Industry in 2050

Abstract

Understanding the effects of climate change is central to assessing the resilience of the agricultural sector in Québec. The dairy industry is vulnerable as climate change alters yields for cattle feed grown on-farm. Québec dairy farmers have adopted various strategies to mitigate greenhouse gas emissions on farms, incorporating sustainable agricultural practices such as improved waste and manure management, and altering cow diets to reduce enteric (digestive) methane production. The last of these practices –– altering cow diets that reduce enteric methane emissions –– is valuable, yet it introduces a tradeoff between emission reduction and climate adaptation. Indeed, diets that reduce methane emissions may require crops that are less resilient to future climate conditions, whereas climate-resilient feed crops may not offer the same methane-reduction benefits. In 2050, Québec dairy farmers may not be able to grow all feed crops on their land to support herd health and milk output, both metrics of feed security. Accordingly, this study assesses the regional feed security of the Québec dairy industry by modelling the impact of crop yield change in two climate scenarios and with three diet compositions in 2050. Results show that in 2050, methane-reducing corn-heavy diets will require more cropland than hay- or soy-based diets, presenting an environmental tradeoff between land use and methane emissions. The analysis reveals high projected intraprovincial variability in feed security, with Eastern Québec predicted to be more feed secure than Southwestern Québec. The importance of a sustainable and self-sufficient dairy industry is increasingly important in the face of climate change. More broadly, this research aims to identify potential approaches for farmers to support future successful dairy operations.

List of Abbreviations

Census Agricultural Region	CAR
Corn-Heavy Diet	CHD
Current Land Area	CLA
Future Land Area	FLA
Greenhouse Gas	GHG
Hay-Heavy Diet	HHD
Intergovernmental Panel on Climate Change	IPCC
Eastern Québec	QCE
Southwestern Québec	QCSW
Representative Concentration Pathway	RCP
Soy-Heavy Diet	SHD

Introduction

Assessing the resilience of key industries is crucial for informed decisions in the face of climate change. Agriculture, including dairy farming, is projected to be one of the industries most impacted by climate change in Canada. The province with the largest dairy industry is Québec, accounting for 46.7% of Canadian dairy farms[1]. Dairy farming accounts for 25% of Québec's agricultural revenue[2]. As the largest agricultural industry in Québec, it also contributes to the fulfillment of Québec's mission of being an economically and culturally self-sustaining province, averting reliance on interprovincial or international trade[2,3]. Existing dairy industry life cycle analyses (LCAs) and agri-food literature primarily acknowledge the climate impacts of greenhouse gas (GHG) emissions from enteric (digestive) fermentation from cattle in the Québec dairy industry[4,5]. Notable research has also been conducted by Cordeiro et al.[6] who assessed the feed security of beef production in the province of Alberta, and Thivierge et al. who simulated future climate impacts at the single-farm scale[7]. Yet, the literature does not account for future climate impacts on dairy production land, a gap which motivates this analysis. This paper seeks to analyze the current and future self-sufficiency of the Québec dairy industry and the potential impacts of climate fluctuations on the sector.

The dairy industry both contributes to and is impacted by climate change. Given this, methane mitigation strategies (MMS) for the beef and dairy industries have been developed by Agriculture and Agri-food Canada[8]. Greenhouse gas emissions of the dairy sector predominantly stem from methane production from enteric fermentation driven by a grass/hay dominant diet[4,5,9]. In accordance, the MMS outlines strategies whereby cattle diets are adapted to reduce enteric fermentation including substituting feed grasses for corn grain and silage or increasing oilseed additives (e.g. soy and canola) in feeds[8]. Several feed crops favored by the MMS are associated with the additional benefit of increased milk yield[9], an important environmental and economic indicator[10]. These suggested feed regimes have already taken effect across the province: Canadian Dairy Farmers estimated that in 2022, 74% of suggested MMS feed practices were already nationally adopted[11]. In Québec, current cow diets consist primarily of hay and are supplemented by corn and soy[4]. The current ratios of corn, grain, silage, and soy are in line with the MMS, though both corn and soy percentages could be further increased under MMS recommendations[11].

In addition to adopting GHG mitigation strategies, dairy farmers must also adapt to climate change-induced crop yield variability, as these fluctuations will affect farmers' ability to produce specific feed crops for their livestock[7]. Québec dairy farmers' heavy reliance on on-farm feed production makes them particularly vulnerable to climate-induced crop yield decreases and variability, with impacts varying in severity across different regions of the

province[4]. In 2022, Québec dairy farms averaged 78 cows and produced about 770,000 litres of milk annually[12], typically operating as non-grazing farms with 92% of cows housed in tie-stalls and cows having an average lifespan of 6 years[4]. According to the Intergovernmental Panel on Climate Change's (IPCC) Representative Concentration Pathway (RCP) 4.5 and 8.5 emissions scenarios, by 2050, Québec will see an increase in average winter temperatures of 2.7 °C and 5.2 °C, and summer increases of 1.9 °C and 2.8 °C, respectively[13]. In these scenarios, the province's agriculture industry will experience nearly 50% longer warm growing seasons and increases in the frequency and intensity of precipitation events[14]. These climate changes will have mixed effects on common dairy cattle feed crops: for instance, soybean, corn, and timothy hay are predicted to benefit from longer frost-free growing seasons[15], whereas canola, wheat, and barley are expected to be negatively impacted[16]. Such variability in crop yields could affect feed security, i.e., the ability of farmers to consistently access adequate feed resources to support herd health and agri-food output (for instance, milk for dairy farmers)[6]. Feed security is an indicator of disruptions in the broader dairy production system. Coupled with climate stress and crop yield variability, the steady rise in Canadian demand for dairy products, including 1-3% increases in cheese, yogurt, and cream consumption in two-year periods[17], requires that cattle feed production meet demand. The combined pressures of changing crop yields and rising production demand may threaten Québec's feed security and ability to maintain a self-sustained dairy industry[18].

As crop yields decline on Québec farms due to climate change, farmers would need to expand their cropland to maintain feed production for their herds. This additional cropland could be diverted from other agricultural uses or newly created from forest or other natural ecosystems. From 2010 to 2015, 65% of land-use change in Southern Canada was forest-to-cropland conversion[19]. These conversions are associated with local biodiversity loss and fragmentation, decreases in natural carbon sequestration, and potential infringement of indigenous sovereignty[20]. Agricultural expansion also occurs at the cost of soil organic carbon sequestration[21]. In Canada and by 2070, unconverted pasture is projected to store 25.3% more carbon dioxide in unconverted pastures than in croplands[22]. As a result, greater cropland conversion resulting from preventing feed insecurity would have major implications for the environments surrounding dairy farms. Decisions about cropland conversion require consideration of the tradeoffs between using the land for dairying or other agricultural operations, and preserving natural ecosystems. Accordingly, this paper aims to assess the feed security of Québec's dairy industry, regionally and provincially, in a future climate-changed world.

Methods

The Model and Calculations

The Model

A feed insecurity model was made in Microsoft Excel. The model allows for the feed insecurity of sub-regions of Québec to be analyzed for two climate scenarios (RCPs 4.5 and 8.5) and with three different diets. The province is divided by Census Agricultural Regions (CARs), as data is most granular at this scale. As of 2024, data on future crop yield projections is available for two subregions of Québec: the East (QCE) and the Southwest (QCSW). As such, CARs are categorized into these two regions, informed by Thivierge and colleagues' map detailing Eastern and Southwestern Québec[7].

Diet Composition

In our analysis, feed insecurity is assessed for three different diets: a hay-heavy diet (HHD) (the current Québec cow diet), a soy-heavy diet (SHD) and a corn-heavy diet (CHD) (the current average US cow diet). These diets were chosen as realistic scenarios in which the fraction of each major feed crop varied. The relative composition of all three diets is shown in Table 1. "Current" refers to the best estimates as of 2024.

Table 1. Diet Composition (%) of the Three Modeled Diets. HHD sourced from Mc Geough et al.[4], CHD from Castillo-Lopez et al.[25], and SHD from Holtshausen et al.[5].

	Hay-Heavy Diet	Corn-Heavy Diet	Soy-Heavy Diet
Corn grain	11.86	16.30	11.56
Corn silage	23.51	36.40	17.11
Hay/Haylage	47.77	9.58	49.12
Soybean	11.66	9.58	13.26

Mc Geough and colleagues'[4] feed composition data is available by cow age category and is divided into 5 life stages. To obtain a lifetime average diet composition, feed components are weighted according to diet composition at each life stage and the duration of each life stage. For each crop, the weighted proportion of the diet (WPD_{crop}) is calculated using Equation 1, where the proportion of the diet (PD) made up for by a given crop during a given life stage is multiplied by the fraction of life (FOL) spent in the given life stage, and where n refers to the life stages 1 to 5:

$$\mathrm{WPD_{crop}} = \sum_{n=1}^{5} \mathrm{PD_{stage\ n}} \times \mathrm{FOL_{stage\ n}}. \quad (1)$$

The average US dairy cattle diet found by Castillo-Lopez et al.[25] is used to simulate a corn-heavy diet. The diet developed by Holtshausen et al.[5] for their work on Québec dairy is used to simulate a soy-heavy diet. All diets are weighted by lifetime consumption.

Mass of Feed Required

The total energy intake of a cow during a given life stage (TEI_{stage}) is the product of the *daily* energy intake during the given life stage (DEI_{stage}) and the duration of the stage (t_{stage}), as shown by Equation 2. The lifetime average daily energy intake of a cow (ADEI), given by Equation 3, is the sum of total energy intake for each life stage, divided by the total lifespan of a cow (t_{total}). The average daily dietary energy ($ADDE_{crop}$) provided by a given crop to a cow throughout its lifetime is the product of ADEI and WPD, as demonstrated by Equation 4. ADDE is found for each crop and each diet.

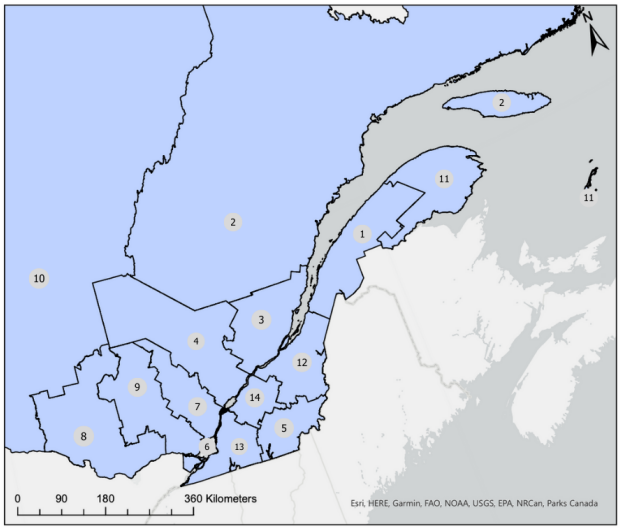
To find the mass of dry matter of a given crop (DM_{crop}) required annually per cow for a given diet, the ADDE of the crop is first converted from units of MJ/day to Mcal/year, and then divided by digestible energy (DE), as shown by Equation 5. The annual crop production required (CPR_{crop}) to meet feed demands is calculated using the dry matter required and the dry matter percentage of the crop ($DM\%_{crop}$), as shown by Equation 6. CPR refers to the mass of wet feed of a given crop required annually per cow, and is calculated for each crop, for each of the three diets. Digestible energy (Mcal/kg) and dry matter percentage values for each crop are derived from the National Research Council's *Nutrient Requirements of Dairy Cattle*[26]:

$$\mathrm{TEI_{stage}\ (MJ)} = \mathrm{DEI_{stage}\ (MJ/day)} \times t_{\mathrm{stage}}\ \mathrm{(days)}, \quad (2)$$

$$\mathrm{ADEI\ (MJ/day)} = \sum \mathrm{TEI_{all\ stages}(MJ)} \,/\, t_{\mathrm{total}}\ \mathrm{(days)}, \quad (3)$$

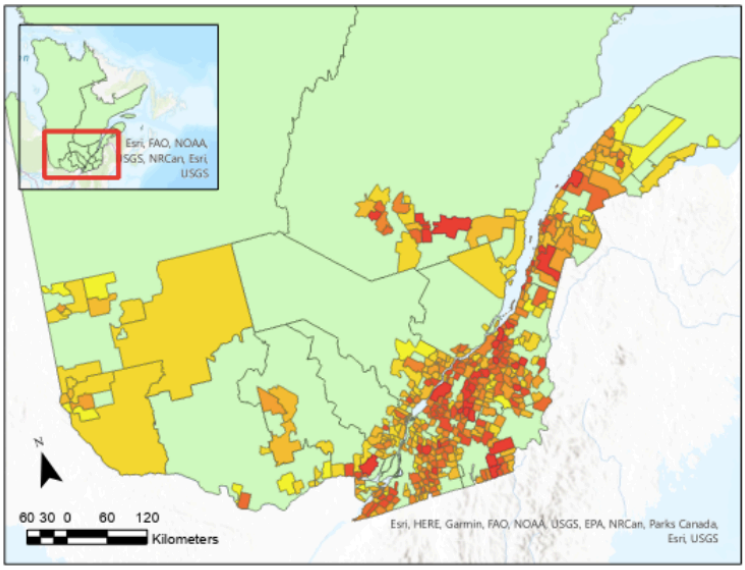
$$\mathrm{ADDE_{crop}\ (MJ/day)} = \mathrm{ADDE\ (MJ/day)} \times \mathrm{WPD}, \quad (4)$$

$$\mathrm{DM_{crop}\ (kg/year/cow)} = \frac{\mathrm{ADDE_{crop}\ (Mcal/year)}}{\mathrm{DE\ (Mcal/kg)}}, \quad (5)$$



1. Bas-Saint-Laurent
2. Saguenay--Lac-Saint-Jean--Côte-Nord
3. Québec
4. Mauricie
5. Estrie
6. Montréal--Laval
7. Lanaudière
8. Outaouais
9. Laurentides
10. Abitibi-Témiscamingue--Nord-du-Québec
11. Gaspésie--Îles-de-la-Madeleine
12. Chaudière-Appalaches
13. Montérégie
14. Centre-du-Québec

(a)



Distribution of Dairy Cows in Québec by Census Consolidated Subdivision

Cattle
70 - 252
253 - 437
438 - 655
656 - 947
948 - 1261
1262 - 1688
1689 - 2210
2211 - 2901
2902 - 3821
3822 - 8414
Quebec CAR borders

(b)

Figure 1. (a) Census Agricultural Regions in Québec. Showing the division of CARs. Data sourced from Statistics Canada[23]. **(b)** Distribution of dairy cows by CCS in Québec. Showing a visual of the distribution of cows by Census Consolidated Subdivisions, a subdivision of CARs. Data sourced from Statistics Canada[23,24].

$$\mathrm{CPR}_{\mathrm{crop}}\ (\mathrm{kg/year/cow}) = \frac{\mathrm{DM}_{\mathrm{crop}}\ (\mathrm{kg/year/cow})}{\mathrm{DM\%}_{\mathrm{crop}}\ (\%)}. \tag{6}$$

Current & Future Number of Cows per CAR

The current number of cows in a given CAR is calculated from Statistics Canada heifer and dairy cattle data[24] by summing the number of heifers and dairy cattle per CAR. According to Alexandratos & Bruinsma[27], it can be assumed that the number of cows in each region increases by 0.3% annually. The number of cows per CAR in 2050 is calculated by Equation 7:

$$\mathrm{Cows}_{2050} = \mathrm{Cows}_{2021} \times (1 + 0.003)^{2050-2021}. \tag{7}$$

Current and Future Land Area Required

Data on the current land area (CLA) used to grow dairy feed crops in Québec is currently unavailable. Literature shows that most dairy producers in Québec grow the required amounts of feed on-farm[4,28]. As feed production mostly occurs on-farm and is currently sufficient for dairy production, we assume that all regions in Québec are currently feed secure. Current crop yield (CCY) data for each CAR is obtained froDiets and Feed Insecuritym the *Institut de la statistique du Québec*[29]. CLA within a given CAR is calculated for each diet scenario. The CLA required to grow the necessary amount of a crop for a given diet within a given CAR (for example the CLA required to grow all the corn grain for the hay-heavy diet in the Bas-St-Laurent CAR) is calculated with Equation 8. The total CLA required to grow feed for a given diet is the sum of the CLA for each individual crop. This total is calculated for each CAR using Equation 9:

$$\mathrm{CLA}_{\mathrm{crop}}\ (\mathrm{ha}) = \frac{\mathrm{Cows}_{2021} \times \mathrm{CPR}_{\mathrm{crop}}\ (\mathrm{kg/year/cow})}{\mathrm{CCY}_{\mathrm{crop}}\ (\mathrm{kg/ha/year})}, \tag{8}$$

$$\mathrm{Total\ CLA}_{\mathrm{CAR}}\ (\mathrm{ha}) = \mathrm{CLA}_{\mathrm{corn\ grain}} + \mathrm{CLA}_{\mathrm{corn\ silage}} + \mathrm{CLA}_{\mathrm{hay}} + \mathrm{CLA}_{\mathrm{soy}}. \tag{9}$$

The total future land area (FLA) required to grow feed for all cows is calculated for each CAR, each diet, and both climate scenarios. Future crop yield changes are sourced from Thivierge et al.[7], which provides data for two greater regions in Québec: Eastern Québec (QCE) and Southwest Québec (QCSW). Future crop yields (FCY) under a given climate scenario are derived from current crop yields (CCY) and the predicted crop yield changes (ΔCY) for the given RCP, as shown by Equation 10. The FLA for a given crop under each RCP is then calculated with Equation 11. Lastly, using Equation 12, the total FLA required to grow feed for a given diet is calculated for each CAR by summing the FLA for each individual crop:

$$\mathrm{FCY}_{\mathrm{RCP,crop}}\ (\mathrm{kg/ha}) = \mathrm{CCY}_{\mathrm{crop}}\ (\mathrm{kg/ha}) + \Delta\mathrm{CY}_{\mathrm{RCP}}\ (\%), \tag{10}$$

$$\mathrm{FLA}_{\mathrm{RCP,crop}}\ (\mathrm{ha}) = \frac{\mathrm{Cows}_{2050} \times \mathrm{CPR}_{\mathrm{crop}}\ (\mathrm{kg/year/cow})}{\mathrm{FCY}_{\mathrm{RCP,crop}}\ (\mathrm{kg/ha})}, \tag{11}$$

$$\mathrm{Total\ FLA}_{\mathrm{CAR}}\ (\mathrm{ha}) = \mathrm{FLA}_{\mathrm{corn\ grain}} + \mathrm{FLA}_{\mathrm{corn\ silage}} + \mathrm{FLA}_{\mathrm{hay}} + \mathrm{FLA}_{\mathrm{soy}}. \tag{12}$$

Feed Insecurity Indicator (FII)

The FII represents the percentage of land area change required to meet feed security. A negative FII indicates a decrease in land required—–land that could be diverted from cow feed production and reclaimed for other purposes such as reforestation—–whereas a positive FII indicates an increase

Table 2. Current Feed Crop Yields and Projected 2050 Feed Crop Yield Changes (%) by CAR. Current yield data obtained from Institut de la Statistique du Québec[29]. Predicted crop yield change data obtained from Thivierge et al.[7]

Census Agricultural Regions (CAR)	Current Yields (kg/ha)				Predicted 2050 Yield Change for RCPs 4.5 and 8 (%)							
	Corn Grain	Corn Silage	Soy	Hay	Corn Grain		Corn Silage		Soy		Hay	
					4.5	8	4.5	8	4.5	8	4.5	8
Bas-Saint-Laurent (QCE)	6500	27150	3110	4790	1.369	-8.22	83.50	76.70	34.78	43.48	15	22.73
Gaspésie – Îles-de-la-Madeleine (QCE)	6500	27150	3110	4790	1.369	-8.22	83.50	76.70	34.78	43.48	15	22.73
Saguenay – Lac-Saint-Jean – Côte-Nord (QCSW)	7680	35190	2050	3420	1.369	-8.22	14.77	16.78	34.78	43.48	15	7.32
Québec (QCSW)	8900	39800	2690	4980	1.369	-8.22	14.77	16.78	34.78	43.48	15	7.32
Mauricie (QCSW)	8900	39800	2690	4980	1.369	-8.22	14.77	16.78	34.78	43.48	15	7.32
Estrie (QCE)	8350	35490	2450	5390	1.369	-8.22	83.50	76.70	34.78	43.48	15	22.73
Lanaudière (QCSW)	9140	48130	3310	5930	1.369	-8.22	14.77	16.78	34.78	43.48	15	7.32
Outaouais (QCSW)	9260	47990	2710	4270	1.369	-8.22	14.77	16.78	34.78	43.48	15	7.32
Laurentides (QCSW)	9260	47990	2710	4270	1.369	-8.22	14.77	16.78	34.78	43.48	15	7.32
Abitibi-Témiscamingue – Nord-du-Québec (QCSW)	7880	19750	1980	2800	1.369	-8.22	14.77	16.78	34.78	43.48	15	7.32
Chaudière-Appalaches (QCE)	7520	32160	2690	5030	1.369	-8.22	83.50	76.70	34.78	43.48	15	22.73
Montérégie (QCE)	10690	40630	3130	6320	1.369	-8.22	83.50	76.70	34.78	43.48	15	22.73
Centre-du-Québec (QCE)	9030	41070	2780	5370	1.369	-8.22	83.50	76.70	34.78	43.48	15	22.73

in land required (conversion of land to cropland):

$$\text{FII} = \frac{\text{Total FLA (ha)}}{\text{Total CLA (ha)}}\% - 100\% \tag{13}$$

Assumptions

Our model functions under the following key assumptions:

[1] All CARs are currently perfectly feed secure.

[2] All feed is grown on-farm.

[3] Future crop yields and milk output are not impacted by improvements in production technology.

[4] The number of cows per CAR will increase by 0.3%/year.

[5] Herds are not impacted by temperature fluctuations (e.g. heat stress).

Analytical Approach

To analyze the sensitivity of feed security to diet changes, the percent of a given crop in the diet composition is increased in 10% increments while decreasing other feed components proportionally.

Interregional variation is assessed to analyze the potential impact of feed insecurity on interprovincial trade. Nuanced analysis of intraprovincial variation and trade projections are outside the scope of this study. As simple metrics of variation, we use standard deviation (SD) and coefficient of variation (CV) from the mean of feed insecurity values across CARs. To note, CAR Montréal–Laval is omitted from all analyses as the region has no dairy farms.

The change in possible milk production is calculated assuming each cow produces 8,000 L of milk annually[30]. All maps were made using the GIS software ArcGIS Pro. Provincial and CAR shapefiles were sourced from Statistics Canada[23].

Results

HHD (Current Québec Diet) Scenario

Under the HHD scenario (current Québec diet), the study finds that all CARs will remain feed secure. The provincial average FII is -6.74% under RCP 4.5 and -10.04% under RCP 8.5. This equates to 39745 ha and 59183 ha, respectively, of reclaimed land area overall in the province. Despite Québec's average negative FII, only 6 of 11 CARs are found to have a negative FII (increased feed security). To meet feed demands, 7 of the 13 included CARs will require an increase in land for feed production by 2050.

The total land area required for feed production by 2050 will decrease in QCE, while increasing in QCSW. QCE is found to have an FII of -9.57% and -14.85% under RCPs 4.5 and 8.5, respectively. This equates to 43170 to 67043 ha of reclaimed land in QCE. Meanwhile, QCSW is found to have an FII of 2.48% and 5.70% under RCPs 4.5 and 8.5, respectively. This equates to 3425 to 7860 additional ha required in QCSW.

Under both RCP 4.5 and 8.5, the Montérégie region exhibits the largest absolute FIIs, while the Lanaudière region exhibits the largest positive FIIs (Table 3).

The Chaudière-Appalaches region sees the largest change in total land area requiring 11619.1 ha (RCP 4.5) to 18143.9 ha (RCP 8.5) less land for feed production by 2050. The Saguenay–Lac-Saint-Jean–Côte-Nord region exhibits the highest increase in total feed cropland area required, 917.1 ha (RCP 4.5) and 2280.3 ha (RCP 8.5).

CHD (Current US Diet) Scenario

In the CHD scenario, Québec is found to remain provincially feed secure under both RCPs, with an FII of -3.01% (8384 ha reclaimed) under RCP 4.5 and -3.18% (8858 ha reclaimed) under RCP 8.5. QCE will require less land to meet feed demands in 2050, with an FII of -6.20% (13552 ha reclaimed) and -6.87% (15003 ha reclaimed) under RCPs 4.5 and 8.5 respec-

Table 3. Québec 2050 Dairy Cow FII and Total Cropland Area Change. FII values expressed as a percentage are given per region (including the provincial, QCE and QCSW values), for each diet, at each RCP. For total cropland area change in Québec, negative numbers represent land reclaimed in 2050, while positive numbers indicate the additional cropland required to meet feed demands in 2050.

Regions	Hay-Heavy Diet RCP 4.5	Hay-Heavy Diet RCP 8.5	Corn-Heavy Diet RCP 4.5	Corn-Heavy Diet RCP 8.5	Soy-Heavy Diet RCP 4.5	Soy-Heavy Diet RCP 8.5
FII for Québec Dairy Cows in 2050						
Bas-Saint-Laurent (QCE)	-8.85%	-13.89%	-5.00%	-5.00%	-11.11%	-16.64%
Gaspésie – Îles-de-la-Madeleine (QCE)	2.17%	5.38%	8.00%	9.00%	-0.15%	2.61%
Saguenay – Lac-Saint-Jean – Côte-Nord (QCSW)	2.73%	5.83%	9.00%	11.00%	0.06%	2.66%
Québec (QCSW)	2.73%	5.83%	9.00%	11.00%	0.06%	2.66%
Mauricie (QCSW)	-9.77%	-15.06%	-7.00%	-7.00%	-11.81%	-17.49%
Estrie (QCE)	4.01%	7.29%	13.00%	15.00%	0.94%	3.67%
Lanaudière (QCSW)	2.73%	6.09%	10.00%	12.00%	0.30%	3.19%
Outaouais (QCSW)	2.73%	6.09%	10.00%	12.00%	0.30%	3.19%
Laurentides (QCSW)	0.77%	3.90%	2.00%	4.00%	-0.78%	2.01%
Abitibi-Témiscamingue – Nord-du-Québec (QCSW)	-8.85%	-13.89%	-5.00%	-5.00%	-11.11%	-16.64%
Chaudière-Appalaches (QCE)	-9.34%	-14.58%	-6.00%	-6.00%	-11.46%	-17.12%
Montérégie (QCE)	-10.38%	-15.72%	-8.00%	-9.00%	-12.16%	-17.88%
Centre-du-Québec (QCE)	-9.46%	-14.95%	-5.00%	-7.00%	-11.47%	-17.33%
Province of Québec	-6.74%	-10.04%	-3.01%	-3.18%	-6.46%	-9.98%
QCE	-9.57%	-14.85%	-6.20%	-6.87%	-9.07%	-14.64%
QCSW	2.48%	5.70%	8.58%	10.20%	2.01%	5.17%
Total Cropland Area Change in Québec in 2050 (ha)						
QC	-39745.2	-59183.0	-8383.5	-8857.8	-38655.3	-59743.9
QCE	-43170.3	-67042.6	-13552.4	-15003.2	-41492.7	-67024.2
QCSW	3425.1	7859.7	5168.9	6145.3	2837.4	7280.3

tively. QCSW will need to increase cropland, with an FII of 8.58% (additional 5169 ha required) and 10.2% (additional 6145 ha required) under RCP 4.5 and 8.5 respectively. Under this diet, Montérégie is the most feed secure CAR, and the least feed secure CAR is Lanaudière for both RCPs (Table 3). Overall, this diet scenario predicts the highest FIIs across both RCPs.

SHD Scenario

The SHD scenario results in an overall feed secure province, with FIIs of -6.46% (38655.3 ha reclaimed) and -9.98% (59743.9 ha reclaimed) under RCPs 4.5 and 8.5 respectively. In RCP 4.5, QCE is found to have a FII of -9.07% (41492.7 ha reclaimed) while QCSW is found to have an FII of 2.01% (additional 2837.4 ha required). In RCP 8.5, the FII is -14.64% (67024.2 ha reclaimed) for QCE and 5.17% (additional 7280.3 ha required) for QCSW. The most feed secure CAR for both RCPs is Montérégie and the least feed secure is Lanaudière (Table 3). This diet scenario leads to the most feed security, producing the lowest FIIs across all included CARs.

Discussion

Feed Security and Diet Composition

The results show a significant impact of diet composition on future feed security. The modeled outcome of the CHD scenario is least feed secure, while the SHD scenario is found to be the most feed secure.

Corn-Heavy Diets

In the CHD scenario, the average FIIs across CARs are 1.92% (σ 8.04%) for RCP 4.5 and 2.69% (σ 9.23%) for RCP 8.5. In contrast, the average FIIs across CARs under the HHD and SHD scenarios for both RCPs are negative, indicating greater feed security on average. Every included CAR presents a higher FII in the CHD scenario than the other diet scenarios.

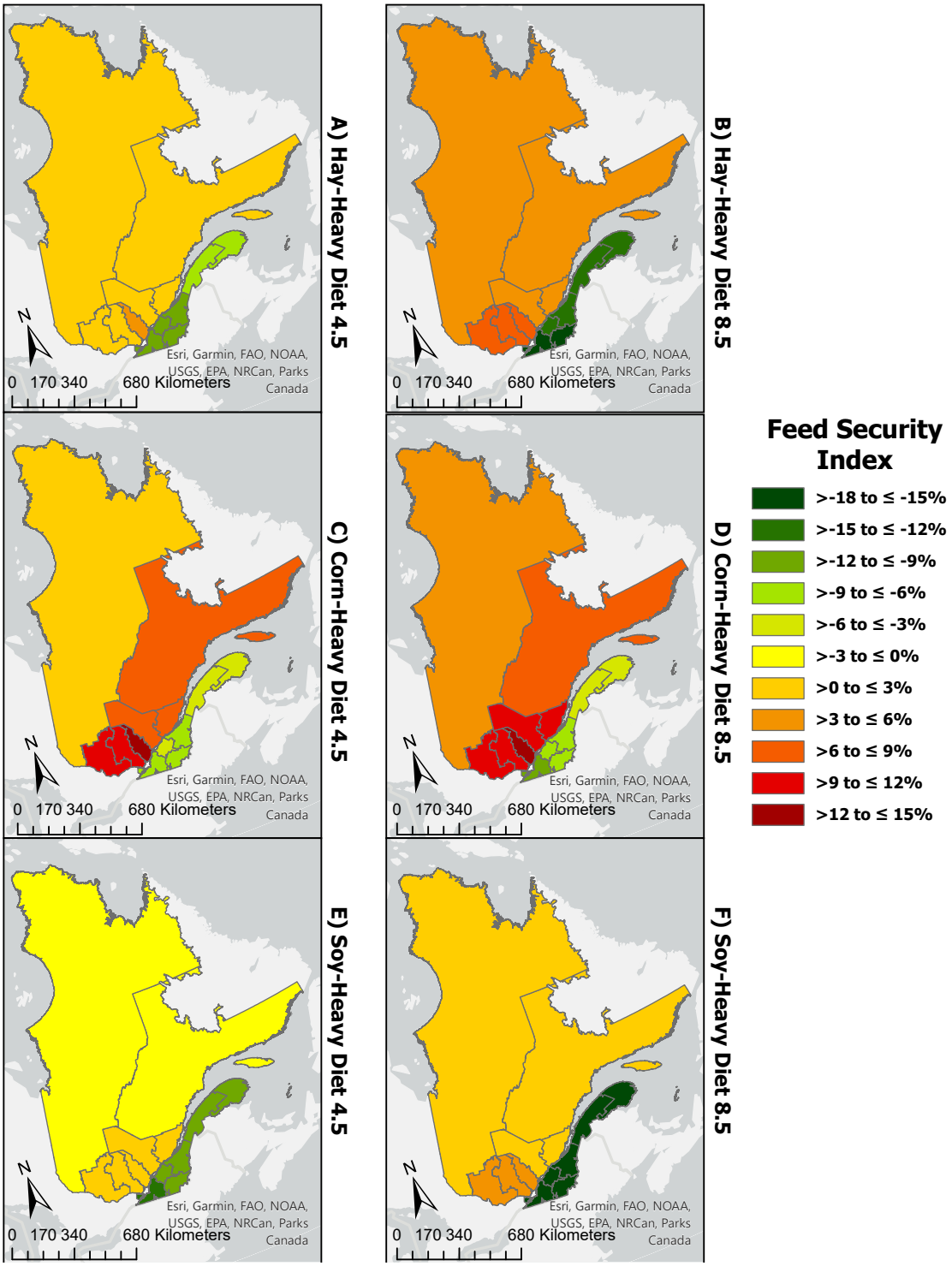
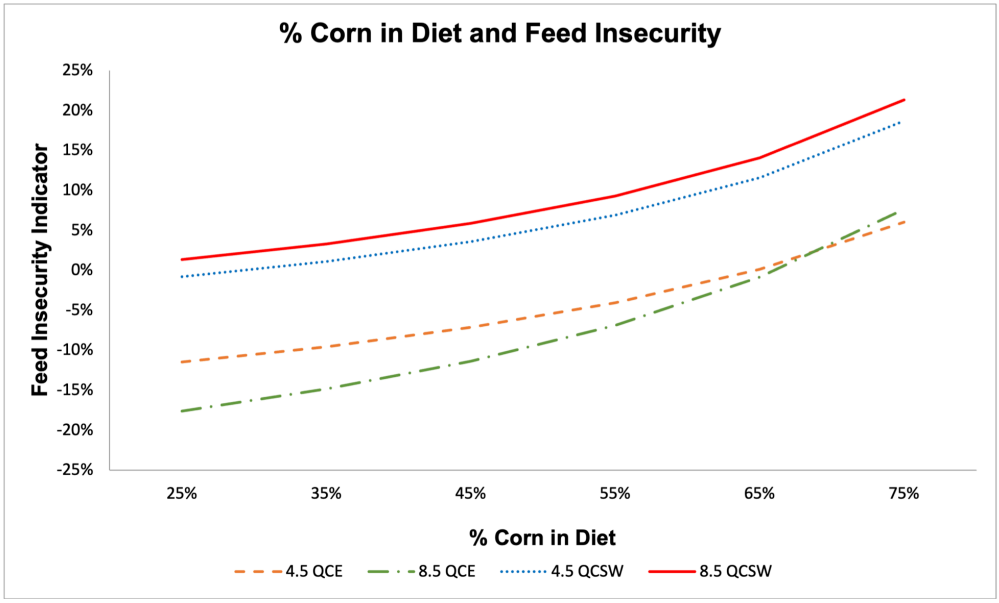


Figure 2. Feed Insecurity Indicator in Québec by CAR. Maps on the left correspond to RCP 4.5, and maps on the right correspond to RCP 8.5.

a)



b)

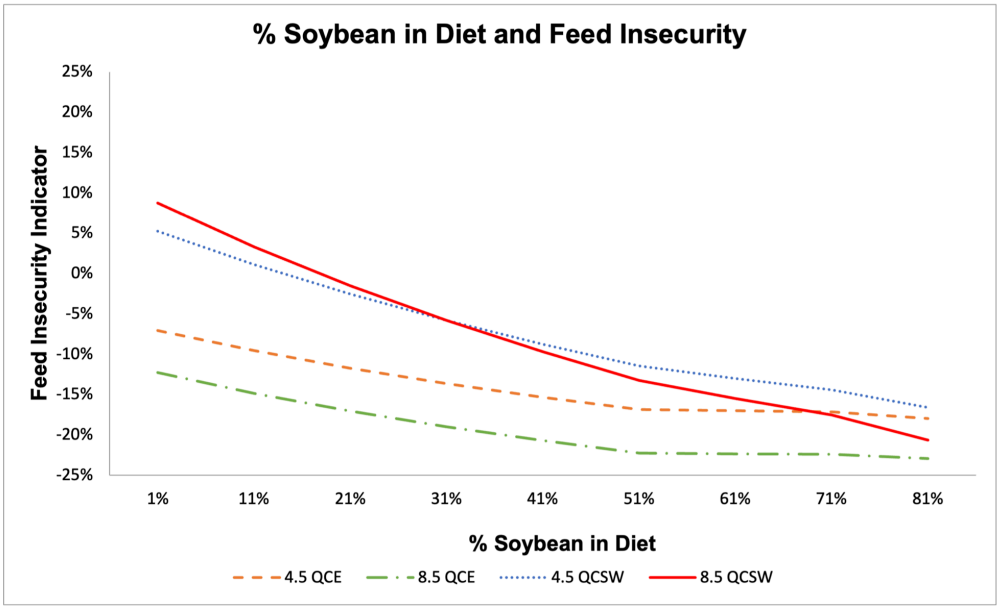


Figure 3. (a) Relationship Between Proportion of Corn (%) in Dairy Cow Diet (Sum of Silage and Grain) and Feed Insecurity. Corn includes both grain and silage. The percentage of corn in the HHD (current Québec dairy cow diet) is 36%. Correlation between percent corn in the diet and FII in QCE and QCSW is $R^2 = 0.958$ for RCP 4.5 QCE, $R^2 = 0.958$ for RCP 8.5 QCE, $R^2 = 0.9432$ for RCP 4.5 QCSW, and $R^2 = 0.9443$ for RCP 8.5 QCSW. (b) Relationship Between Proportion of Soybean (%) in Dairy Cow Diet and Feed Insecurity. The percentage of soybeans in the HHD (current Québec dairy cow diet) is 11.66%. Correlation between percent soybean in the diet and FII in QCE and QCSW is $R^2 = 0.9082$ for RCP 4.5 QCE, $R^2 = 0.8921$ for RCP 8.5 QCE, $R^2 = 0.9722$ for RCP 4.5 QCSW, and $R^2 = 0.9765$ for RCP 8.5 QCSW.

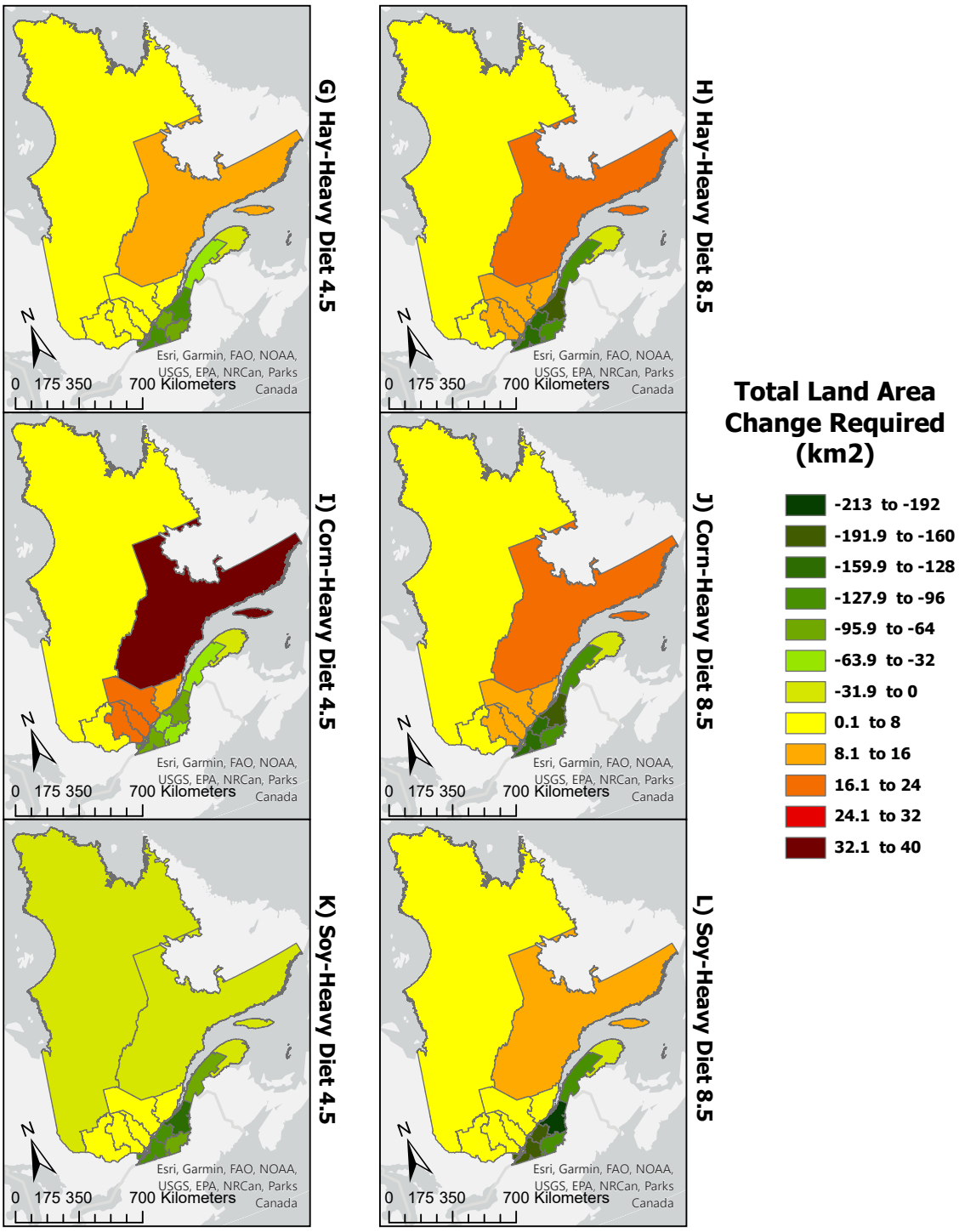


Figure 4. Total land area change required per CAR (km^2). Maps show the total land area change required in Québec by CARs. Maps on the left correspond to RCP 4.5, and maps on the right correspond to RCP 8.5.

This demonstrates that a corn-heavy diet renders regions at higher risk of feed insecurity than other diets. This is primarily a result of the projected change in corn grain yields, which increase only modestly under RCP 4.5 and decrease under RCP 8.5 (Table 2). Other crop yields considered in this study, including corn silage, are projected to increase under both RCPs. While corn grain comprises 44% of the corn intake in the CHD (corn silage comprising 56%), its baseline yield is considerably lower than that of corn silage, requiring 330% more land on average in Québec (Table 2). Thus, the decrease in corn grain yield counteracts the increases in corn silage yield. This leads us to conclude that the proportion of corn grain in the diet is among the largest risk factors of projected feed insecurity in Québec.

Despite this diet scenario presenting the highest feed insecurity indicators across all included CARs, Québec dairy farmers will experience a province-wide increase in feed security. However, due to high intraprovincial variation in feed security, not all feed will be able to be produced on-farm.

The choice to increase corn feed demonstrates a tradeoff in the environmental sustainability of the dairy industry. Hassanat et al.[9] found that when fed a high-starch, corn-based diet, there was a quadratic reduction in methane production in Québec dairy cows. This study also found that methane reductions only occurred when corn comprised >50% of the diet by weight, suggesting a critical proportion higher than the current Québec diet which contains 36% corn[9]. Lastly, Hassanat et al.[9] revealed increased milk yield under a corn-heavy diet: When comparing a no-corn diet to a 50% corn diet, the study found a 10% increase in milk yield in the cattle fed a corn-heavy diet. Similarly, Guyader et al.[31] found an 8.7% increase in milk yield in Québec dairy cattle when barley feed was replaced with corn. This indicates a tradeoff between methane-reducing strategies and land use expansion for agriculture.

A comparative LCA is beyond the scope of this work. However, we note that if cropland increases in proximity to existing dairy farms, most of this land use change will occur in Southern Québec (Figure 1). Changes in land use in Southern Québec have implications on the mitigation of greenhouse gas emissions, as most terrestrial carbon (forest biomass and soil organic carbon) is found in this part of the province[32]. Thus, increases in the dairy feed cropland to provide a CHD could disturb Québec's most dense forest and soil carbon areas.

Soy-Heavy Diets

In the SHD scenario, the average feed insecurity indicators across included CARs for RCP 4.5 and 8.5 are -5.26% (σ 6.05%) and -6.39% (σ 10.41%), respectively. Every CAR presents a lower feed insecurity indicator in the SHD scenario than the other diet scenarios.

This demonstrates that a soy-heavy diet renders regions less at risk of feed insecurity than other diets. This is largely because soy yield is projected to increase significantly (Table 2). An additional factor is the decreased corn consumption in the SBD, which also predicts improved feed security.

A SHD may present a similar greenhouse gas-related environmental tradeoff. Soybean meal, particularly when compared to canola meal, has been shown to increase enteric fermentation (methane emissions) and nitrous oxide emissions from manure[5,33]. Canola is the most available and equivalent protein soy replacement for Canadian dairy cattle; canola yields are projected to decline in Canada in the face of climate change[16]. This leads us to predict that replacing soybean with canola to reduce methane emissions would increase feed insecurity and drive larger land use demands. However, further research must be done to examine the impact of climate change on feed security to determine the tradeoff of GHG emissions and land use.

Feed Security and Intraprovincial Regional Variation

According to Statistics Canada[34], total cropland area in Québec is 1,942,491 ha. Thus, the increase in dairy cattle feed security seen across these various diet and RCP scenarios represents a 2-3% reduction in total required cropland in the province (8383.5 to 59743.9 ha) by 2050.

The impacts of climate change are non-uniform in Québec; therefore, feed security is predicted to vary regionally. Under all scenarios, QCE is more feed secure than QCSW. This is because corn silage and hay yields are projected to increase dramatically more in QCE than QCSW (Table 2) due to greater increases in mean temperature, annual precipitation, and changes in planting dates for various crops[7].

Feed Security and Intraprovincial Trade

Currently, the literature indicates that feed for dairy cattle in Québec is grown on-farm[4,28]. However, climate change is projected to increase geographical variation in crop yields across the province. This increased variation, which would project some CARs with a feed surplus and others with a feed deficit, will likely drive intraprovincial trade. Intraprovincial trade, rather than sourcing feed from elsewhere within Canada, is particularly likely given the province's desire to maintain a self-sustaining dairy industry.

Standard deviation (SD) and coefficient of variation (CV) are interpreted as simple metrics of intraprovincial variation. The SD for all diet and RCP scenarios is high, with the CV ranging from 192% to 418%. These values indicate substantial variation across CARs, leading us to predict greater trade of feed crops in 2050.

Increased transportation for trade will lead to increased GHG emissions from the dairy sector. These emissions will be particularly significant if the variation occurs largely east-to-west, implying long-haul transports from QCE to QCSW.

The CHD scenario demonstrates the highest CV values: 418% and 342.90% under RCP 4.5 and 8.5, respectively. This demonstrates an increased trade demand under this diet scenario compared to a HHD or SHD, highlighting an additional environmental tradeoff of the CHD: while it may reduce enteric fermentation, it would increase both transportation and land use requirements.

Possible Milk Production

As Québec becomes more feed secure, increased crop production could feed more cattle. If the current HHD remains, in 2050 Québec will be able to produce enough crops on existing land to feed an additional 38,906 dairy cows under RCP 4.5 and 54,026 dairy cows under RCP 8.5. This equates to approximately 61,000,000 and 436,000,000 additional liters of milk annually, 0.84% and 1.17% of current Québec milk production[17].

Conversely, regions that experience feed insecurity will be unable to support the current number of dairy cattle without expanding land use. QCSW, which is predicted to face feed insecurity in 2050, will be unable to support 2,635 of the region's projected dairy cattle if land use remains stagnant. This equates to a loss of 21,082,847 liters of milk annually from the QCSW region.

Trade or land use expansion must account for this loss to prevent adverse economic and social outcomes of dairy production decline.

Greater Implications

Although this study focused on the tradeoffs between feed security and methane reduction, it is important to note the broader, additional implications of different diets in dairying. Such implications include impacts on biodiversity, pesticide use, water use, and water pollution. Most notably, the production of soy- and corn-heavy diets requires significantly greater pesticide use[35].

Additionally, the study highlights the possibility for feed strategies to reduce methane emissions from the dairy industry. The IPCC's policy recommendations describe the potential of consumer plant-based diets for climate change mitigation and adaptation[36]. Future work would benefit from assessing the impact of these diets on methane emissions.

Limitations

The model's assumptions inherently restrict the accuracy of the outcomes. To ensure model consistency, both crop yield and milk yield are assumed fixed. For model simplicity, we assume that crop yield would not be impacted by factors other than climate change. In tandem, milk yield is also kept constant across RCPs, despite the likelihood that heat stress will reduce milk yield in a warming future[37]. Further LCAs which aim to quantify future GHG emissions should consider accounting for not only yield and land-use change, but also increased transport emissions.

The limited availability and geographical scope of data also constrains our results. Two RCP pathways, 2.6 and 6, are not employed in this study as crop projections aligned with these scenarios are currently not found in literature. Thivierge and colleagues' predicted temperature and yield increases do not consider all possible extreme climate weather events[7]. Projected crop yields also compartmentalized data only into two regions: QCE and QCSW. Because of this, QCE and QCSW data was coarsely applied to their respective CARs.

Conclusion

The feed supply of the Québec dairy industry will endure various impacts of climate change. This study predicts that as crop yields fluctuate with warming and increased precipitation, Québec will become more feed secure, requiring 2-3% less cropland in 2050 to feed the projected dairy cattle than at the time of writing. Meanwhile, regional variation in crop production is projected to increase. If land use remains stagnant, Eastern Québec will produce a feed surplus while Southwestern Québec will be in feed deficit. This intraprovincial variability will drive feed crop trade, fundamentally changing the current on-farm feed production system and driving an increase in industry GHG emissions from transportation. This study also revealed that diet has a profound impact on feed security. Corn-heavy diets are projected to render Québec more feed insecure than other diets would, requiring 374% to 573% more land for crop production than soy- or hay-heavy diets, as soy and hay yields will be favorably impacted by climate change. Corn-heavy diets are also projected to drive the largest regional crop yield variation, increasing trade requirements and thereby transport emissions. While corn-based diets are recommended for methane reduction, they create an environmental tradeoff by requiring greater land use. Further research on Québec dairy's environmental sustainability should consider predicted land use changes. In the context of Québec's dedication to methane reduction strategies, further investigations of the GHG tradeoffs of various diets would be pertinent to determine the most sustainable diets.

Acknowledgement

This study would not have been possible without Dr. Graham MacDonald, Associate Professor of Geography at McGill University, whose insight and expertise in agricultural systems and modelling was invaluable.

References

1. Canadian Dairy Commision. Number of Farms with Shipments of Milk on August 1st. https : / / agriculture . canada . ca / en / market - information - system / rp / index - eng . cfm ? action = gR & r = 220&signature=2BB1B27ED46C50FE77D7BC142B757B69&pdctc= &pTpl=1#wb-cont (2023).
2. Agriculture and Agri-Food Canada. Overview of Canada's agriculture and agri-food sector (2023); https://agriculture.canada.ca/en/sector/overview
3. Ministère de l'Agriculture, des Pêcheries et de l'Alimentation (MAPAQ). *Agir, pour une agriculture durable — Plan 2020-2030 [Acting, for sustainable agriculture – 2020-2030 plan]* (2020).
4. Mc Geough, E. et al. Life-cycle assessment of greenhouse gas emissions from dairy production in Eastern Canada: A case study. *J. Dairy Sci.* **95,** 5164–5175 (2012). https://doi.org/10.3168/jds.2011-5229
5. Holtshausen, L., Benchaar, C., Kröbel, R. & Beauchemin, K. A. Canola meal versus soybean meal as protein supplements in the diets of lactating dairy cows affects the greenhouse gas intensity of milk. *Animals* **11,** 1636 (2021). https://doi.org/10.3390/ani11061636
6. Cordeiro, M. et al. Assessing feed security for beef production within livestock-intensive regions. *Agric. Syst.* **196,** 103348 (2022). https://doi.org/10.1016/j.agsy.2021.103348
7. Thivierge, M.-N. et al. Projected impact of future climate conditions on the agronomic and environmental performance of Canadian dairy farms. *Agric. Syst.* **157,** 241–257 (2017). https://doi.org/10.1016/j.agsy.2017.07.003
8. Agriculture and Agri-Food Canada. Reducing Methane Emissions from Livestock (2019); https://agriculture.canada.ca/en/science/story-agricultural-science/scientific-achievements-agriculture/reducing-methane-emissions-livestock
9. Hassanat, F. et al. Replacing alfalfa silage with corn silage in dairy cow diets: Effects on enteric methane production, ruminal fermentation, digestion, N balance, and milk production. *J. Dairy Sci.* **96,** 4553–4567 (2013). https://doi.org/10.3168/jds.2012-6480
10. Lorenz, H., Reinsch, T., Hess, S. & Taube, F. Is low-input dairy farming more climate friendly? A meta-analysis of the carbon footprints of different production systems. *J. Clean. Prod.* **211,** 161–170 (2019). https://doi.org/10.1016/j.jclepro.2018.11.113
11. Dairy Farmers of Canada. *Net Zero by 2050: Best Management Practices Guide to Mitigate Emissions on Dairy Farms* (2023).
12. Les Producteurs de Lait du Québec. *Supply Management and Collective Milk Marketing* (2023).
13. Environment Canada. *Climate data and scenarios for Canada: Synthesis of recent observation and modelling results* (2016).
14. Li, G. et al. Indices of Canada's future climate for general and agricultural adaptation applications. *Clim. Chang.* **148,** 249–263 (2018). https://doi.org/10.1007/s10584-018-2199-x
15. Alberti-Dufort, A. et al. Quebec; Chapter 2 in *Canada in a Changing Climate: Regional Perspectives Report*. (eds Warren, F., Lulham, N., Dupuis, D. & Lemmen, D.). Government of Canada (2022).

16. Qian, B. et al. Simulated canola yield responses to climate change and adaptation in Canada. *Agron. J.* **110,** 133–146 (2018). https://doi.org/10.2134/agronj2017.02.0076

17. Les Producteurs de Lait du Québec. *Rapport Annuel 2022 [Annual Report 2022]* (2022).

18. Ouranos. *Vers l'adaptation. Synthèse des connaissances sur les changements climatiques au Québec. Édition 2015 [Toward adaptation. Synthesis of knowledge on climate change in Québec. 2015 Edition]* (2015).

19. Government of Canada. Land-use change (2022); https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/land-use-change.html

20. KC, K. B. et al. Opportunities and trade-offs for expanding agriculture in Canada's North: an ecosystem service perspective. *Facets* **6,** 1728–1752 (2021). https://doi.org/10.1139/facets-2020-0097

21. Wei, X., Shao, M., Gale, W. & Li, L. Global pattern of soil carbon losses due to the conversion of forests to agricultural land. *Sci. Rep.* **4,** 4062 (2014). https://doi.org/10.1038/srep04062

22. Jiang, R. et al. Impacts of land use conversions on soil organic carbon in a warming-induced agricultural frontier in Northern Ontario, Canada under historical and future climate. *J. Clean. Prod.* **404,** 136902 (2023). https://doi.org/10.1016/j.jclepro.2023.136902

23. Statistics Canada. 2021 Census –- Boundary Files (2021); https://www12.statcan.gc.ca/census-recensement/2021/geo/sip-pis/boundary-limites/index2021-eng.cfm?year=21

24. Statistics Canada. Table 32-10-0370-01 Cattle inventory on farms, Census of Agriculture, 2021. https://doi.org/10.25318/3210037001-eng (2022).

25. Castillo-Lopez, E. et al. Performance of dairy cows fed silage and grain produced from second-generation insect-protected (Bacillus thuringiensis) corn (MON 89034), compared with parental line corn or reference corn. *J. Dairy Sci.* **97,** 3832–3837 (2014). https://doi.org/10.3168/jds.2014-7894

26. National Research Council. in *Nutrient Requirements of Dairy Cattle: Seventh Revised Edition, 2001.* 281–289 (The National Academies Press, 2001). https://doi.org/10.17226/9825.

27. Alexandratos, N. & Bruinsma, J. *World agriculture towards 2030/2050: the 2012 revision* ESA Working paper No. 12-03. Rome, FAO (2012). https://doi.org/10.22004/ag.econ.288998

28. Guyader, J., Little, S., Kröbel, R., Benchaar, C. & Beauchemin, K. *Carbon footprint of dairy production systems in Québec: Barley versus corn silage* in Abstracts of the 2017 American Dairy Science Association® Annual Meeting, June 25–28, 2017, Pittsburgh, PA. *J. Dairy Sci.* **100** Suppl. 2 (2017). https://doi.org/10.1016/S0022-0302(20)30813-4

29. Institut de la statistique du Québec. Area of field crops, yield per hectare and production, by combined administrative regions, Québec, 2007-2023 (in French only). https://statistique.quebec.ca/en/produit/tableau/3786 (2023).

30. Fédération des producteurs de lait du Québec. *The journey of MILK from the farm to your table* (2014).

31. Guyader, J., Little, S., Kröbel, R., Benchaar, C. & Beauchemin, K. A. Comparison of greenhouse gas emissions from corn-and barley-based dairy production systems in Eastern Canada. *Agric. Syst.* **152,** 38–46 (2017). https://doi.org/10.1016/j.agsy.2016.12.002

32. Sothe, C. et al. Large soil carbon storage in terrestrial ecosystems of Canada. *Glob. Biogeochem. Cycles* **36,** e2021GB007213 (2022). https://doi.org/10.1029/2021GB007213

33. Benchaar, C., Hassanat, F., Beauchemin, K., Gislon, G. & Ouellet, D. Diet supplementation with canola meal improves milk production, reduces enteric methane emissions, and shifts nitrogen excretion from urine to feces in dairy cows. *J. Dairy Sci.* **104,** 9645–9663 (2021). https://doi.org/10.3168/jds.2020-20053

34. Statistics Canada. Over one-third of dairy cows were reported in Québec (2011); https://www150.statcan.gc.ca/n1/pub/95-640-x/2011001/p1/prov/prov-24-eng.htm

35. Nehring, R. et al. *Pesticide Use in US Agriculture: 21 Selected Crops, 1960-2008. EIB-124* U.S. Department of Agriculture, Economic Research Service (2014).

36. Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2023).

37. West, J. W. Effects of heat-stress on production in dairy cattle. *J. Dairy Sci.* **86,** 2131–2144 (2003). https://doi.org/10.3168/jds.S0022-0302(03)73803-X

Submitted: 1/12/2024
Accepted: 04/03/2025
Published: 03/31/2025

Research Article

[1]Department of Psychology, McGill University, Montréal, QC, Canada

Keywords

Mice, Pain, Grimacing, Lateralization, Asymmetry

Email Correspondence

elodie.nickner@mail.mcgill.ca

https://doi.org/10.26443/msurj.v20i1.218

Elodie Nickner[1]

Asymmetry of Pain-Induced Facial Grimacing

Abstract

Pain has two main components: the sensory-discriminative (the quality, intensity and location of pain) and the motivational-affective (the emotional aversiveness of pain reflective of suffering)[1,2]. A plethora of translational preclinical and clinical measures for the sensory-discriminative component exist (e.g., von Frey, cold pressor). However, few existing measures capture the more elusive motivational-affective component, and those that do are hampered as they are not translatable across species. Post-lesion evaluation of facial grimacing of emotion-related areas of the brain suggests that the Mouse Grimace Scale is reflective of the motivational-affective component[3]. Facial expressions of emotion (e.g., fear, anger) are lateralized such that the left side of the face exhibits facial expressions more strongly than the right side[4,5]. Comparing pain-induced facial grimacing to facial expressions of emotion is one way to determine which component of the pain experience is most captured by the Mouse Grimace Scale. We hypothesized that grimacing would be lateralized to the left side of the face. Examining lateralization of pain-induced facial grimacing is novel to pain research. We examined the asymmetry of pain-induced facial grimacing in CD-1 mice using inflammatory, neuropathic, and reflexive pain. And we found that pain is expressed predominantly on the right side of the face, contrary to other emotions. Our findings have important implications for the measurement of pain, as characterized by suffering, in non-verbal populations and for application in veterinary care settings.

Introduction

Pain plays an important adaptive function, despite being an aversive experience. Melzack and Casey (1968) identified two components of the pain experience, the sensory-discriminative (SD) and motivational-affective (MA) components[6]. The SD component reflects the quality, intensity, and location of pain, while the MA component reflects the aversive emotional suffering of pain. The International Association for the Study of Pain recognizes pain as characterized by both a "sensory and emotional experience"[7]. Despite recognition of both the SD and MA components, numerous measures of the SD component exist, such as von Frey, which measures mechanical withdrawal thresholds, while measures capturing the MA component remain ill-defined. Evidence suggesting that the Mouse Grimace Scale (MGS) may capture the more elusive MA component of pain stems from lesion studies showing that lesions to limbic regions in the brain associated with emotion processing, such as the amygdala, led to the attenuation of pain-induced facial grimacing[3]. Considering that suffering and the aversive characteristics of pain are ubiquitous, it is essential to identify well-characterized translational methods that objectively measure these components in pain research.

Pain Assessment Scales

Self-report questionnaires (e.g., visual analogue scales (VAS), numerical rating scales (NRS), verbal rating scales (VRS)) are often used to capture the subjective pain experience in humans. To reduce self-report bias present in subjective measures, Paul Ekman developed the Facial Action Coding System (FACS): an objective measure capturing changes in facial musculature in response to stimuli[8,9]. By the late 20th century, pain was finally recognized in infants, and the Neonatal Facial Coding System (NFCS) was adapted from the FACS, enabling pain assessment in infants[10].

The Mouse Grimace Scale

While the FACS and NFCS are limited to use in humans, cross-species generalizability of pain captured through facial expressions emerged when researchers collaborated in the development of the MGS[11]. Four features of the MGS distinguish it from all other measures in algesiometry—the measurement of pain sensitivity as a response to stimuli (e.g., pressure, heat)[12]. First, well-established measures of pain capture localized pain (e.g., von Frey); however, most pain syndromes are characterized by *diffuse* pain (e.g., visceral pain)[12]. The MGS captures pain broadcasted via facial features, i.e., the more commonly reported diffuse type of pain. The second is the $duration$ of pain[12]. Longstanding measures of pain (e.g., hot-plate, tail-flick test) are limited in the sense that they capture pain of short duration (e.g., seconds to minutes). While such measures were pioneering in the field, they do not capture the most clinically relevant type of pain, that is chronic pain lasting months to years. However, the MGS is a measure that captures facial expressions of pain over extended periods of time. The third feature is pain $modality$; more specifically, experimenter-evoked versus spontaneous pain[12]. A major disadvantage of the established measures of pain is that they capture experimenter-evoked pain rather than spontaneous pain. Spontaneous pain is of greater clinical significance as it relates more closely with chronic pain. The MGS mends this issue by capturing facial expressions of pain in response to spontaneous pain across various modalities (mechanical, thermal, chemical, or electrical) by relying on the broadcast of pain via facial features. The fourth and final component is *outcome measures* (e.g., subjective vs objective)[12]. Established pain assessment scales capturing the MA component are limited to use in humans due to their requirement for written or verbal communication through subjective questionnaires. The MGS is an objective measure that captures the MA component via the external broadcast of facial features, permitting translation of this model across species. We argue that the MGS is one such tool permitting the preclinical study of the more clinically significant pain—chronic pain characterized as diffuse, spontaneous, and longer lasting.

Pain: a Sensation and an Emotion

Pain encompasses both sensory and affective features, however, the more researched of the two remains the sensory component. Recognition of the sensory component is ubiquitous, as demonstrated by pain being recognized as the fifth vital sign in the field of medicine[13]. However, the affective component of pain, characterized by subjective suffering, remains ill-captured across most species. Most established preclinical models aim to capture affect measure behaviours (e.g., tail suspension and conditional place preference) as proxies for the emotionality of pain[14]. As research using rodents is characterized by a lack of verbal assessments, we cannot be certain about the emotional component of their pain. Because the MA component of pain is characterized by suffering, arguably the more relevant aspect of the pain experience, it is crucial to clearly define translational measures, such as the MGS, to directly capture this component.

Lateralization of Facial Expressions of Emotion

Darwin first noted the asymmetricity of facial expressions of emotion[15]. Prior studies have shown that facial expressions of emotion (e.g., fear, anger) are predominantly expressed on the left side of the face[16]. In both human and non-human samples, left asymmetry of emotion was found in facial expressions via third-party interpretation, and in brain areas involved in their expression[4]. External left-bias for emotional expression is thus also reflected in the internal neural circuitry governing emotion expression. Innateness of left asymmetry in facial expressions of emotion was shown in research examining intensity ratings of spontaneous versus posed expressions of emotion[17]. Examination of spontaneous versus posed facial expressions of positively and negatively valenced emotions found a stronger left-side display bias for spontaneous versus posed emotions for both happy and sad emotion types. To our knowledge, whether pain is expressed asymmetrically remains unknown. We examined whether spontaneous pain-induced facial grimacing would be lateralized, such that the expression of pain would be stronger on the left or on the right side of the face, using the MGS.

Current Study

Despite current knowledge on facial grimacing, it remains unclear to what degree the MGS captures the sensory or the emotional aspects of pain. With similar objective grimace scales developed in humans, we have reason to believe that the MGS captures the MA component; therefore, we examined whether pain-induced facial grimacing is lateralized. We hypothesized that grimacing would be lateralized to the left side of the face like facial expressions of emotion. Mice underwent one of five different pain models: zymosan (ZYM), complete Freund's Adjuvant (CFA), acetic acid (AA), carrageenan (CARR), and spared nerve injury (SNI), across three different pain types: neuropathic, inflammatory, and reflexive pain. The SD component of pain was captured using von Frey as an established tool for the assessment of mechanical withdrawal thresholds, with fibres ranging from 0.008 grams to 100 grams of force. Lateralization of MGS scores and von Frey withdrawal thresholds were compared. Using a within-subjects design, mice facial grimacing was coded according to the 5 facial action units (AUs) comprising the MGS coding system.

Materials and Methods

Video Capture and Pain Model Induction

Overview

The appropriate research review board, Institutional Animal Care and Use Committee (IACUC), reviewed and approved this research. Adult male and female CD-1 mice on a 12/12-hour light-dark cycle underwent 30-minute baseline video recordings prior to any experimental manipulation. Subsequently, mice were assigned to an experimental group, AA, CFA, ZYM, SNI, or CARR, and injected with the assay. All injections were administered into the hind paw and counterbalanced between left and right, with the exception of AA, which was administered via intraperitoneal injection (IP). IPs of AA were completed by scruffing the mouse and using 1-mL syringes with 26-gauge needle tips to inject into the intraperitoneal cavity. The concentrations and volume of AA for each mouse were determined by mouse weight using a standard dosage formula. CFA, ZYM, and CARR injections were administered into the pad of the hind paw using 1-mL syringes with 26-gauge needle tips. Similarly, the concentration and volume of the pain assay were determined by mouse weight using a standard dosage formula. All injections were completed by a rotating group of trained individuals. SNI surgeries were completed by the same research assistant, ensuring standardization of the procedure. All animals were anesthetized using isoflurane before the surgery. Skin and muscle incisions were made before ligating the terminal tibial and peroneal branches of the sciatic nerve using a non-resorbable silk suture, followed by a resection of a small segment of these two branches[18]. The sural branch remained intact. In sham controls, no sciatic nerve branches were ligated or resected, they were only exposed. The skin incision was closed in all animals using absorbable silk suture material. Post-injection video recordings were taken following a window of time appropriate for each agent: AA, immediately after injection[19]; CFA, 2hrs post-injection[20]; ZYM 30-45 minutes post injection[21]; SNI, 4 days post injection[22]; and CARR, 3 hours post-injection[20].

Video Capture

High-definition (HD) video recordings of mice displaying facial grimacing and other relevant pain behaviours were stored on SD cards prior to being uploaded to a server for storage and evaluation. A group of undergraduate students were responsible for taking these videos. Mice were individually placed in single stainless steel and Plexiglas observation cubicles ($9 \times 5 \times 5$ cm high) with HD digital video cameras positioned perpendicular to the cubicle and facing the mouse.

Image Generation and Standard MGS Scoring

Still images were taken from both baseline and post-model induction videos at 3-minute intervals. 7 images were generated for each 3-minute time interval corresponding to front-facing (unaltered), left and right composite (composite), left and right side (hemiface), and left and right profile view (profile) picture orientations. Still images were cropped so that the body position of the mice was not within the frame. Blinded and randomized scoring of still images was performed to obtain an average baseline and post-induction MGS score for each of the 5 AUs (orbitals, whiskers, ears, cheek and nose bulge), as well as an overall score. A single research assistant was tasked with blindly scoring the stills using the MGS. Analysis of the AA videos was completed separately by counting the number and duration of abdominal constrictions (writhes). Writhing behaviour was defined as repeated constriction of the abdominal muscles, followed by elongation of the body and hind-leg extension[23]. Five research assistants quantified writhing behaviour by manually counting the number of writhes exhibited by the mice.

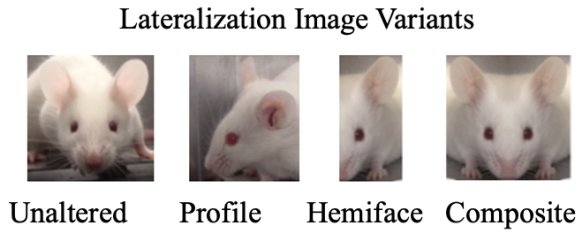


Figure 1. Lateralization Image Variants: 4 image types of mice faces were collected from video recordings; unaltered front facing images, both left and right side-profile images, both left and right hemiface images, and both left and right composite images made from two identical left or right hemiface images mirrored to make a symmetrical face.

Generation of Composite Images

Symmetrical (unaltered) still images of the mice's faces were sampled approximately every 3 minutes and were cropped down the y-axis and mirrored to create left-left and right-right facial chimeras. The cropped still images used for chimeras were also used without mirroring, using only the left or right side of the face (hemiface).

Scoring Using the Mouse Grimace Scale

The MGS applies a 5-AU scoring system that includes the degree of orbital squinting, nose and cheek bulge, as well as ear and whisker position change[3]. Rating of AUs is on a scale from 0 to 2 depending on the degree of expression (0 = AU not present, 1 = moderate presence of the AU, 2 = severe presence of the AU). The AUs that were not clearly determined were not scored.

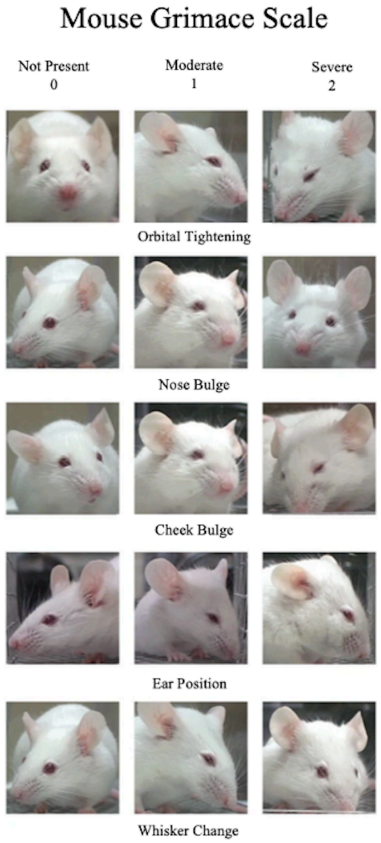


Figure 2. Mouse Grimace Scale: The Mouse Grimace Scale assesses the absence, moderate or severe presence of five action units: (1) orbital tightening, (2) nose bulge, (3) cheek bulge, (4) ear position, and (5) whisker change.

Capturing the Sensory Component Using von Frey

von Frey (vF) was used to capture the sensory component of pain to serve as a methodological control condition against the MGS—believed to capture the affective component of pain. All vF measurements were taken by the same research assistant who remained blind to the conditions. Mice were habituated for 30 minutes before vF measurements were taken.

Results

All analyses were conducted using Statistical Package for Social Sciences (SPSS), and graphs were generated using Prism version 7.04. To test our hypothesis that pain-induced facial grimacing is lateralized to the left side of the face, we analyzed raw MGS scores and compared them to the average MGS scores for each AU. We included a total of n=164 mice in our study, and each data point in the graphs represents one mouse. This analysis was done for each picture orientation: front facing (unaltered), left and right composite (chimera), left and right side only (hemiface), and left and right profile view (profile). Baseline (pre-pain induction) MGS scores were subtracted from post (post-pain induction) MGS scores to obtain a difference/delta score (D MGS). We analyzed these scores using a two-tailed one-sample t-test comparing mean scores to zero. Scores greater than zero indicate a right bias, whereas scores below zero indicate a left bias. The red error bars represent the standard error of the means, whereas the grey areas signify the means. We observed significant lateralization of pain-induced facial grimacing, with pain being expressed predominantly on the right side of the face across all studied angles of the face: composite ($t(162)=2.757$, $p=0.0065$); hemiface ($t(162)=3.685$, $p=0.0003$); profile ($t(162)=8.486$, $p<0.0001$); average ($t(162)=6.682$, $p<0.0001$).

Delta MGS Scores Separated by Side View

A two-tailed one-sample t-test was conducted comparing the delta MGS scores for each AU to zero. We observed a significant effect of orbitals $t(161)=2.418$, $p=0.0167$, nose bulge $t(161)=3.836$, $p=0.0002$, and cheek bulge $t(161)=2.912$, $p=0.0041$. However, the ears $t(161)=0.6841$, $p=0.495$ and whiskers $t(154)=1.666$, $p=0.0978$ AU scores were not significantly different from 0.

Delta MGS Scores Separated by Action Unit

A two-tailed one-sample t-test was conducted comparing delta MGS scores depending on the side of injection of the pain assay. There was a significant right-bias in MGS scores for pain-induced facial grimacing, which occurred regardless of the side of pain injection. Pain injection on the right ($t(73)=4.370$, $p<0.0001$), left ($t(72)=4.236$, $p<0.0001$), and non-lateral IP ($t(15)=2.831$, $p=0.0127$) all caused stronger pain-induced facial grimacing on the right side of the face.

Delta MGS Scores Relative to Site of Injection of Pain Assay

A two-tailed one-sample t-test was conducted comparing the delta MGS scores for each pain assay. We observed a significant effect of SNI ($t(29)=4.088$, $p=0.0003$), AA ($t(15)=2.831$, $p=0.0127$), ZYM ($t(32)=3.075$, $p=0.0043$), and CFA ($t(59)=3.860$, $p=0.0003$). The delta MGS scores for CARR ($t(23)=1.407$, $p=0.172$) were not significantly different from 0.

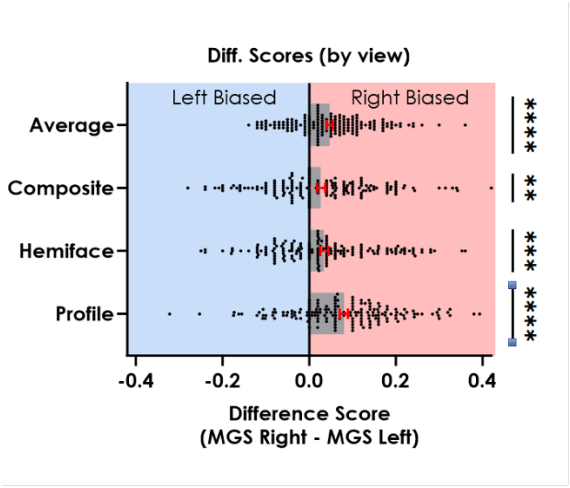


Figure 3. MGS Difference Scores Shown by Face Side: Pain-induced facial grimacing was expressed more strongly to the right side of the face from every facial angle. Composite $t(162)=2.757$, $p<0.01$, $M=0.02727$, $SD=0.1262$; hemiface $t(162)=3.685$, $p<0.001$, $M=0.03461$, $SD=0.1199$; profile $t(162)=8.486$, p<0.0001, $M=0.07984$, $SD=0.1201$; average $t(162)=6.682$, $p<0.0001$, $M=0.04730$, $SD=0.09037$. **$p\leq0.01$, ***$p\leq0.001$, ****$p\leq0.0001$.

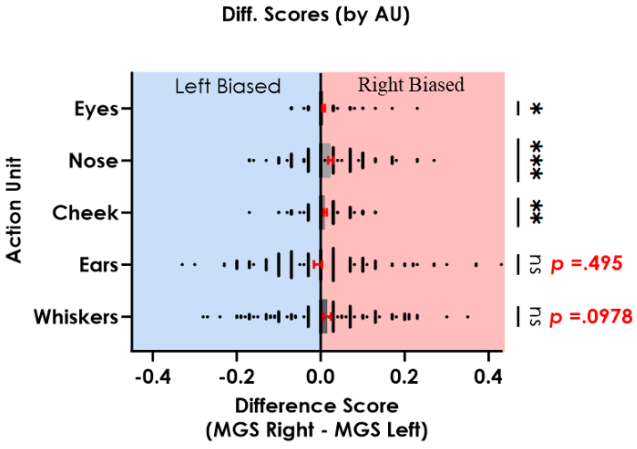


Figure 4. MGS Difference Scores by Action Unit: Mice showed a bias in pain-induced facial grimacing to the right in the eye ($t(161)=2418$, $p<0.05$, $M=0.006790$, $SD=0.03575$), nose ($t(161)=3.836$, $p<0.001$, $M=0.01056$, $SD=0.1217$), and cheek ($t(161)=2.912$, $p<0.01$,M= 0.01056, $SD=0.04613$). The ears ($t(161)=0.6841$, p = ns, $M=0.006543$, $SD=0.1217$) and whiskers ($t(154)=1.666$, $p<0.05$, $M=0.01523$, $SD=0.1138$), although trending towards right-biased lateralization, did not reach significance, which indicates differential lateralization of grimacing action units. *$p\leq0.05$, **$p\leq0.01$, ***$p\leq0.001$, ns = nonsignificant.

Delta MGS Scores Relative to Pain Assay

An independent samples t-test, assuming equal variance, was conducted to investigate whether von Frey withdrawal thresholds were lateralized. The side of injury predicted ipsilateral and contralateral withdrawal thresholds congruently, such that a left side injury predicted lower withdrawal thresholds on the ipsilateral (left) side ($t(109)=1.509$, $p=0.134$) and high withdrawal thresholds on the contralateral (right) side ($t(109)=-1.036$, $p=0.302$). This result suggests that mechanical hypersensitivity is not lateralized.

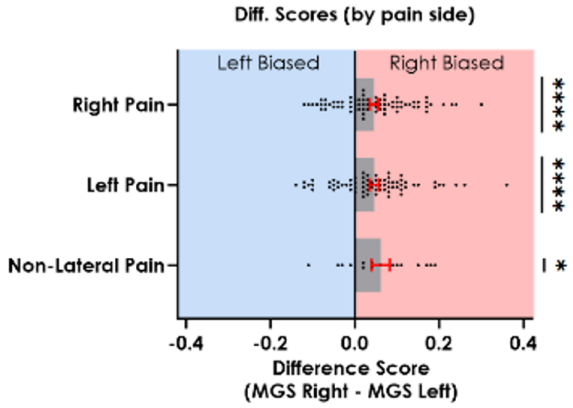


Figure 5. MGS Difference Scores Shown by Pain Location: Pain-induced facial grimacing was predominantly expressed on the right side of the face regardless of pain location; right pain ($t(73)=4.370$, $p<0.0001$, $M=0.04554$, $SD=0.08964$), left pain ($t(72)=4.236$, $p<0.0001$, $M=0.04603$, $SD=0.09284$), non-lateral pain ($t(15)=2.831$, $p<0.05$, $M=0.06125$, $SD=0.08655$). *$p\leq0.05$, ****$p\leq0.0001$.

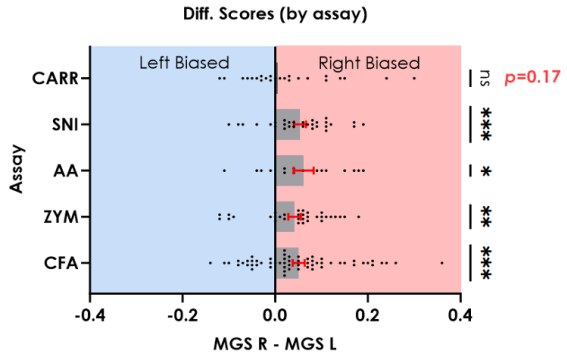


Figure 6. MGS Difference Scores by Inflammatory Assay: Mice showed a strong lateralization of facial grimacing response to inflammatory assays: SNI ($t(29)=4.088$, $p<0.001$, $M=0.05333$, $SD=0.07145$), AA ($t(15)=2.831$, $p<0.05$, $M=0.06125$, $SD=0.08655$), ZYM ($t(32)=3.075$, $p<0.01$, $M=0.04182$, $SD=0.07812$), and CFA ($t(59)=3.860$, $p<0.001$, $M=0.05067$, $SD=0.1017$), but no lateralization was observed in response to CARR ($t(23)=1.407$, p = ns, $M=0.02958$, $SD=0.1030$). *$p\leq0.05$, **$p\leq0.01$, ***$p\leq0.001$, ns = nonsignificant.

Discussion

MGS scores for pain-induced facial grimacing were greater for the right side of the face than the left. von Frey withdrawal thresholds were congruent with the side of pain and thus not lateralized. Specifically lower withdrawal thresholds were associated with the pain side whereas higher withdrawal thresholds were associated with the no-pain side. We hypothesized that pain-induced facial grimacing would be lateralized to the left side of the face like other facial expressions of emotion, such as fear. Our results did not support this hypothesis, suggesting that pain does not behave like other facial expressions of emotion. These findings are novel to the field of pain and hold major pre-clinical significance.

Importantly, our results suggest that the MGS captures the more elusive MA component, as suggested by the incongruency between the lateralization of pain-induced facial grimacing and the absence of lateralization of von Frey withdrawal thresholds, which are thought to capture the SD component of pain. This incongruency suggests that the pain measurements MGS and von Frey capture different components of the pain experience, notably the MA and the SD respectively. Characterization of the MGS as the first pre-clinical measure of the MA component of pain opens important avenues in the field, particularly as it pertains to pain assessment in underrepresented,

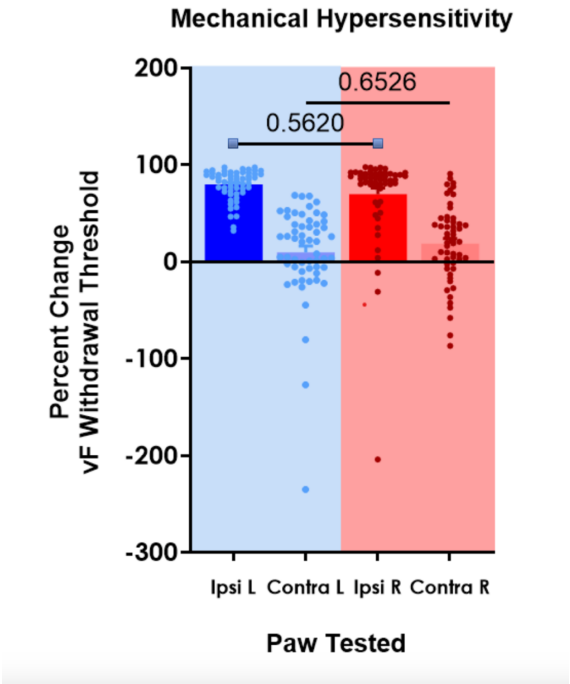


Figure 7. von Frey Withdrawal Thresholds and Site of Injury: Variance in withdrawal thresholds were congruent with the site of injury. Left side injury would predict lower von Frey withdrawal thresholds on the ipsilateral side and higher von Frey withdrawal thresholds on the contralateral side, suggesting that mechanical hypersensitivity is not lateralized. Ipsi-L versus ipsi-R ($t(218)$=1.303, p < ns, $M_{ipsi\text{-}L}$=79.64, $M_{ipsi\text{-}R}$=69.69); contra-L vs. contra-R ($t(218)$=-1.160, p < ns, $M_{contra\text{-}L}$=9.604, $M_{contra\text{-}R}$=18.47).

non-verbal populations. We argue that the MGS, due to its ability to capture characteristics resembling that of chronic pain, should serve as a golden standard in pre-clinical and clinical pain research. Utilizing the MGS in pre-clinical pain assessments brings the field one step closer to demystifying the nature of chronic pain.

One possible explanation as to why pain and emotion expression are not congruently lateralized is in their evolutionary underpinnings. Emotions communicate socially meaningful information (e.g., social norms)[24], while the physiological experience of pain communicates vital information reflective of survival (e.g., the reflex to remove one's hand from a hot stove), thus reflecting the evolutionary basis for segregated information processing pathways[25]. Our results reflect this notion as the communication of pain and emotion, indicated by facial grimacing, is differentially lateralized.

Limitations

Despite the success of the MGS in detecting facial expressions related to pain, such tools might not fully capture the emotional and psychological dimensions of pain that are present in humans. Emotional suffering in pain is multifaceted, involving not just sensory inputs but cognitive assessments, personal history, and cultural context, all of which contribute to how pain is experienced. Therefore, while the MGS and similar tools can provide valuable data on the sensory aspects of pain (e.g., the physical distress indicated by grimacing), they may not account for the entire spectrum of pain's emotional impact in humans. This study underscores the importance of refining measures of the motivational-affective component in both animal and human models to improve our understanding of pain.

Conclusion

In summary, we found that pain-induced facial grimacing is lateralized to the right side of the face, contrary to expressions of emotion. Segregation of the emotion and pain pathways may reflect evolutionary adaptations, as evidenced by differential lateralization of facial expressions for pain and emotion. This finding highlights the MGS as the only pre-clinical measure translatable across species that effectively captures the MA component of chronic pain. This has critical implications for the treatment of suffering and the adversity of chronic pain.

Acknowledgments

I would like to acknowledge Dr. Jeffrey Mogil and PhD candidate Alicia Zumbusch for their estimable guidance throughout the course of this research project. Their invaluable support and assistance served a pivotal role in the execution of this project. I would like to thank my fellow undergraduate students, Gabriel Firanescu, Dana Harrel, Sijie Xu, as well as research assistant 2, Mrs. Susana Sotocinal, for their tremendous help throughout the video-taking and scoring process.

Statement of Contribution

I had the pleasure of working on a research project, part of a larger whole, under Dr. Jeffrey Mogil and PhD candidate Alicia Zumbusch. Collection of MGS videos along with scoring was assisted by fellow undergraduates and stills were collected on a weekly basis. Alicia Zumbusch helped with the collection of background literature, developed the project methodology and provided guidance towards relevant sources. As the eldest undergraduate, I trained my fellow undergraduates on wet-lab tasks and video still extraction. I conducted the data analysis with the help of Dr. Jeffrey Mogil and Alicia Zumbusch and interpretation of the results was subsequently approved by the latter.

References

1. Melzack, R. & Wall, P. D. Pain Mechanisms: A New Theory: A gate control system modulates sensory input from the skin before it evokes pain perception and response. *Science* **150,** 971–979 (1965). https://doi.org/10.1126/science.150.3699.971
2. Talbot, K., Madden, V., Jones, S. & Moseley, G. The sensory and affective components of pain: are they differentially modifiable dimensions or inseparable aspects of a unitary experience? A systematic review. *Br. J. Anaesth.* **123,** e263–e272 (2019). https://doi.org/10.1016/j.bja.2019.03.033
3. Langford, D. J. et al. Coding of facial expressions of pain in the laboratory mouse. *Nat. Methods* **7,** 447–449 (2010). https://doi.org/10.1038/nmeth.1455
4. Lindell, A. Lateralization of the expression of facial emotion in humans. *Prog. Brain Res.* **238,** 249–270 (2018). https://doi.org/10.1016/bs.pbr.2018.06.005
5. Mandal, M. K. & Ambady, N. Laterality of Facial Expressions of Emotion: Universal and Culture-Specific Influences. *Behav. Neurol.* **15,** 23–34 (2004). https://doi.org/10.1155/2004/786529
6. R., M. & L., C. in *The Skin Senses* (eds Kenshalo, D. R.) 423–443 (Charles C Thomas, 1968).

7. Raja, S. N. et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *Pain* **161,** 1976–1982 (2020). https://doi.org/10.1097/j.pain.0000000000001939

8. Prkachin, K. M. The consistency of facial expressions of pain: a comparison across modalities. *Pain* **51,** 297–306 (1992). https://doi.org/10.1016/0304-3959(92)90213-U

9. Ekman, P. & Friesen, W. V. Measuring facial movement. *Environ. Psychol. Nonverbal Behav.* **1,** 56–75 (1976). https://link.springer.com/article/10.1007/BF01115465

10. Rodkey, E. N. & Riddell, R. P. The infancy of infant pain research: the experimental origins of infant pain denial. *J. Pain* **14,** 338–350 (2013). https://doi.org/10.1016/j.jpain.2012.12.017

11. Langford, D. J. et al. Coding of facial expressions of pain in the laboratory mouse. *Nat. Methods* **7,** 447–449 (2010). https://doi.org/10.1038/nmeth.1455

12. Mogil, J. S. The history of pain measurement in humans and animals. *Front. Pain Res.* **3,** 1031058 (2022). https://doi.org/10.3389/fpain.2022.1031058

13. Scher, C., Meador, L., Van Cleave, J. H. & Reid, M. C. Moving beyond pain as the fifth vital sign and patient satisfaction scores to improve pain care in the 21st century. *Pain Manag. Nurs.* **19,** 125–129 (2018). https://doi.org/10.1016/j.pmn.2017.10.010

14. Salehpour, F., Mahmoudi, J., Sadigh-Eteghad, S. & Cassano, P. in *Photobiomodulation in the Brain* (eds Hamblin, M. R. & Huang, Y.-Y.) 189–205 (Elsevier, 2019). https://doi.org/10.1016/B978-0-12-815305-5.00044-0

15. Darwin, C. *The Expression of the Emotions in Man and Animals* (John Murray, 1872).

16. Lindell, A. K. The silent social/emotional signals in left and right cheek poses: A literature review. *Laterality: Asymmetries of Body, Brain and Cognition* **18,** 612–624 (2013). https://doi.org/10.1080/1357650X.2012.737330

17. Dopson, W. G., Beckwith, B. E., Tucker, D. M. & Bullard-Bates, P. C. Asymmetry of facial expression in spontaneous emotion. *Cortex* **20,** 243–251 (1984). https://doi.org/10.1016/s0010-9452(84)80041-6

18. Shields, S. D., Eckert III, W. A. & Basbaum, A. I. Spared nerve injury model of neuropathic pain in the mouse: a behavioral and anatomic analysis. *J. Pain* **4,** 465–470 (2003). https://doi.org/10.1067/s1526-5900(03)00781-8

19. Gawade, S. P. Acetic acid induced painful endogenous infliction in writhing test on mice. *J. Pharmacol. Pharmacother.* **3,** 348 (2012).

20. Ren, K. & Dubner, R. Inflammatory models of pain and hyperalgesia. *Inst. Lab. Anim. Res. J.* **40,** 111–118 (1999). https://doi.org/10.1093/ilar.40.3.111

21. Meller, S. & Gebhart, G. Intraplantar zymosan as a reliable, quantifiable model of thermal and mechanical hyperalgesia in the rat. *Eur. J. Pain* **1,** 43–52 (1997). https://doi.org/10.1016/s1090-3801(97)90052-5

22. Menorca, R. M., Fussell, T. S. & Elfar, J. C. Nerve physiology: mechanisms of injury and recovery. *Hand Clin.* **29,** 317–330 (2013). https://doi.org/10.1016/j.hcl.2013.04.002

23. Dai, G. et al. Synergistic interaction between matrine and paracetamol in the acetic acid writhing test in mice. *Eur. J. Pharmacol.* **895,** 173869 (2021). https://doi.org/10.1016/j.ejphar.2021.173869

24. Van Kleef, G. A., Cheshin, A., Fischer, A. H. & Schneider, I. K. The social nature of emotions. *Front. Psychol.* **7,** 896 (2016). https://doi.org/10.3389/fpsyg.2016.00896

25. Garland, E. L. Pain processing in the human nervous system: a selective review of nociceptive and biobehavioral pathways. *Prim. Care Clin. Off. Pract.* **39,** 561–571 (2012). https://doi.org/10.1016/j.pop.2012.06.013

Submitted: 05/06/2023
Accepted: 01/05/2025
Published: 03/31/2025

Review Article

[1]Department of Physics, McGill University, Montréal, QC, Canada

Keywords

Quantum gravity, String theory, Cosmology, Quantum field theory, Algebraic topology

Email Correspondence

jeffrey.morais@mail.mcgill.ca

https://doi.org/10.26443/msurj.v20i1.337

Jeffrey Morais[1]

Conflicts with de Sitter Vacua in Superstring Theory

Abstract

Models of our universe lack consistency at different energy scales, so we require a theory with ultraviolet (UV) completion such as string theory. A suitable candidate to model our universe in this framework is de Sitter space, a spacetime which expands and has positive curvature. When describing the expansion of this space, however, one computes the wrong sign for the cosmological constant that would not allow for an expanding universe. This motivates one to consider corrections from a quantum theory to reproduce the correct positive sign for the cosmological constant. The conditions that cause this incorrect sign are known as the *swampland criteria*, and prevent de Sitter space from being realized in a consistent manner at different energy scales. We look at a framework to avoid the swampland restriction in a UV-complete theory by considering de Sitter space resulting from compactifications of type IIB superstring theory. In particular, we demonstrate that the definitions of particles in an expanding UV-incomplete theory leads to inconsistencies in the definition of the de Sitter vacuum states. Furthermore, we review previous attempts to prevent these inconsistencies by constructing coherent states that expand and have the desired de Sitter isometries over supersymmetric Minkowski space. These states add quantum corrections to the metric operator, resulting in the cosmological constant carrying the correct sign. Therefore, the de Sitter space can be used in a UV-complete theory to model our universe.

Introduction

It is of great interest in theoretical physics to develop a complete theory of our universe that remains consistent across all energy scales. Current theories exist at very small energy scales (and thus large length scales), like general relativity, or at very high energy scales (and small length scales), such as quantum field theory (QFT). Typically, a theory is defined up to a certain energy scale — *a cutoff* — and above it the equations of motion do not apply. Extending a theory beyond this cutoff to make it well-defined at arbitrarily high energy scales is termed an *ultraviolet (UV) completion*. In a UV-complete theory, the features at high energy scales imply those at low energy scales, and vice versa. In this sense, there is a mixing between the degrees of freedom of theories at high energy scales (ultraviolet) and low energy scales (infrared), known as *UV/IR mixing*. A theory consistent across all energy scales must therefore be UV-complete and exhibit UV/IR mixing. Since such a theory is local in nature, it requires a local quantum field theory description[1]. So far, two theories that have accurately predicted natural phenomena are QFT and general relativity. However, neither exhibits UV/IR mixing, and since they are defined for different energy scales, they do not share the same features; hence, they do not form a complete theory. The attempt to reconcile a quantum description of particles with a classical theory of gravity is known as the problem of *quantum gravity*.

One candidate theory of quantum gravity is string theory, a UV-complete field theory. Although string theory (hence superstring theory as well) is self-consistent, meaning alternate computation methods yield the same results, scientists are still unable to construct a model that reproduces the required properties of nature. Furthermore, many theories that do reproduce these properties, such as QFT or general relativity, reveal flaws under closer inspection, including divergent observables and geometries. In this paper, we are concerned with the *cosmological constant*. This constant determines the universe's rate of expansion: a positive sign corresponds to an expanding universe while a negative sign (or vanishing constant) does not. Under a theory with an expanding spacetime, classical computations do not yield the right sign due to conditions known as the *swampland criteria*. This motivates the introduction of quantum corrections from a quantum theory to bypass these classical constraints, allowing theories to fall outside of the *swampland*. By using type IIB superstring theory, we construct a UV-complete model of our current universe — which contains IR/UV mixing — and avoid the divergence issues of other theories.

To begin, we live in a 3+1D universe (3 spatial dimensions, 1 time dimension) where the curvature of spacetime is nearly flat. Type IIB string theory, however, is consistent in 9+1D, which is six more dimensions than our universe. How then can we recover our lower-dimensional universe from this higher-dimensional theory? The process is called *compactification*[⋆]. The extra six dimensions form a *compact internal space* that, roughly speaking, we take to be small. This process is similar to the construction of a Riemann sphere in complex analysis, whereby a point from infinity is brought to the complex plane to form the compact Riemann sphere. A visualization of the compactification of a torus is shown in Figure 1.

In our case, we split the 9+1D theory into a 3+1D piece (our universe) and a 6D piece (the internal space), and recover our universe via compactification of the internal space. Now that we can recover our universe from a higher dimensional one, the next step is to select which universe model to use. One positively curved and expanding candidate model is 4D de Sitter space dS_4, a vacuum solution to the Einstein field equations. In it, the vacuum states (not to be confused with vacuum *solutions*) share the symmetries of the space and allow for the computation of observables one would measure in a laboratory. One such observable is the *dark energy*, directly related to the spacetime's cosmological constant. As previously mentioned, this constant is crucial as it determines the universe's expansion rate, with a positive constant expected for expanding space.

To construct de Sitter vacua within a UV-complete theory, one method is

[⋆]Compactification is a process in general topology where we take a topological space or manifold, usually one of the extra/internal higher dimensions of the theory, and make it into a compact space[2]. The physics definition extends to taking this compact space to vanish in the limit that the parameter which modulates its size vanishes (such as taking the radius of an n-sphere S^n to vanish: $R \to 0$).

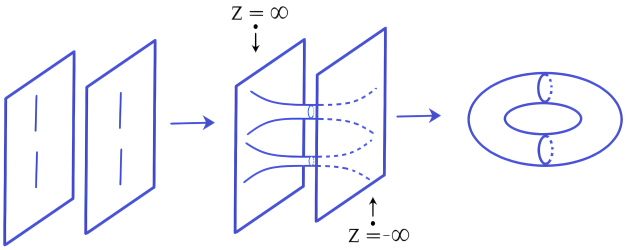


Figure 1. Constructing a torus via compactification. On the left are two parallel 2-dimensional planes with two defined branch cuts over each plane. We connect the branch cuts to get the topology in the middle. Thereafter, we compactify both planes by bringing in points from infinity. The resulting shape is topologically equivalent (diffeomorphic) to a torus.

the KKLT scenario[3], in which a metastable de Sitter state is constructed by uplifting an AdS state via some branes[†] in the presence of a warped geometry in type II string theory. However, it has been shown that this scenario cannot have a well-defined effective field theory as it meets the *swampland criteria*[5], and hence cannot be realized as a consistent theory of gravity. Additionally, in the KKLT regime, the de Sitter conjecture adds no new information in the weak coupling regions where vacua like KKLT are claimed to lie and is violated by the Higgs potential[6]. Another attempt is the Bunch-Davies vacuum, described via a Fourier decomposition of modes over a static patch of de Sitter space[7]. However, because these modes diverge in amplitude at the boundary of the patch, one cannot control the ground state, which may evolve into an excited state. Finally, through compactifications of supergravity, one can construct a de Sitter space (and hence state) subject to certain *classical conditions*[8], under which the components of the metric on the space diverge after compactification. This means that there cannot exist consistent vacuum solutions (i.e. de Sitter space) from compactifications of string theory, leading again to the so-called *swampland* scenario.

As explained, although classically these solutions are forbidden, we can use a quantum theory to obtain quantum corrections that will bypass these conditions and let de Sitter space occur outside the swampland region. We can then obtain a consistent theory of our expanding universe, using de Sitter space, that admits the correct sign for the cosmological constant. Instead of examining vacuum states in expanding geometries like de Sitter space, we instead turn to *Glauber-Sudarshan* states[9]. These generalized coherent states contain all the degrees of freedom of the fields present in string theory and share the isometries of de Sitter space over supersymmetric Minkowski space. Although supersymmetric Minkowski space is flat and does not expand, Glauber-Sudarshan states are expanding and possess the required de Sitter isometries, meaning that the combination remains a relevant candidate to model our universe.

In this paper, we present that in an expanding space — within a UV-incomplete theory — the definitions of particles (and hence their associated vacuum states) become no longer well-defined due to time-dependent frequencies. Because we need a proper definition of vacua states to compute observables in a quantum theory, we require the use of an *alternate* formulation in which the space is taken to be static and the vacua dynamic (which shares the symmetries of the original dynamic space). We look at this alternate formulation by reviewing how coherent states over supersymmetric Minkowski space may be used as a description of vacua in an expanding spacetime. These states add quantum corrections to the space's metric operator, resulting in a positive sign for the cosmological constant. This means that after compactification, the space expands without divergent metric components, making de Sitter space a candidate to model our universe in a UV-complete theory.

[†] In certain topological versions of string theory, *branes* can be viewed as subspaces of Calabi-Yau manifolds[4], a special type of Ricci flat manifold.

de Sitter Vacuum States

We begin by examining how fields behave on expanding geometries in UV-incomplete theories. This is relevant as these fields are exact excitations of the vacuum states in de Sitter space used to calculate the cosmological constant. Specifically, we show that it is impossible to formulate an effective action in an accelerating spacetime because the fields develop time-dependent frequencies. This means we cannot integrate out higher energy modes of fields, and thus *cannot define a theory* at a fixed energy scale. This shows that vacuum solutions of the Einstein field equations in general relativity, including de Sitter space, cannot have consistent descriptions of matter (massive particles) when the solutions are expanding spacetimes. The simplest case we can consider is scalar bosons in an expanding geometry, a review of which is given in Mukhanov, V. & Winitzki, S. (2007)[10], where it is shown that scalar fields develop time-dependent effective mass terms in the Lagrangian, meaning particles have no inherent description of mass. We are interested in cases of higher spin—such as vector bosons or spinor fermion—which could potentially have different statistics or behaviours. To demonstrate the definition of particles breaks down in a dynamic background, we will consider particles with different statistics to scalar bosons which have higher spin: spinor fermions. It is noted that as well as it has different complex structure for its corresponding bundle) see how the definition of particles breaks down.

First, the de Sitter space we are working with is curved and expands to model our universe spacetime. To define fields on equal footing within UV-complete theories over a spacetime manifold, we define them as sections of a *fibre bundle*[11]. A fibre bundle is a non-trivial collection of manifolds that generalizes the notion of product spaces between said manifolds. Since the latter can change at different points, using fibre bundles allows us to define more general field configurations for use in UV-complete theories. A fibre bundle of two manifolds is composed of a *total space* and a *base space*, along with a projection mapping between the two. A fibre bundle can be seen as a prescription of information selection: the base space tells us *how* to select information from the total space. A useful property of fibre bundles is that locally, the space resembles a typical product space. This allows one to work in locally flat spacetime – by defining a local trivialization that uses locally flat charts of $\mathbb{R}^n$. These manifolds are connected by a projective mapping π, which acts on the elements of the total space known as *fibres*. Naturally, the total space is the collection of all the fibres. Furthermore, different fibre bundles are *associated* together if a morphism can be defined between them, such that we can construct one bundle from the other. For example, consider a *spacetime* bundle where the total space is Euclidean space $\mathbb{R}^3$, and the base space is $\mathbb{R}^1$ representing a time dimension. One bundle that can be associated to this spacetime bundle is a *vector* bundle which has the same base space of $\mathbb{R}^1$ as the spacetime bundle, however its fibres are instead vector spaces over the fibres of the spacetime bundle. It is on these vector field fibres that fields (and hence particles) are defined. In the context of type IIB superstring theory in 9+1D, we work with a more general spacetime bundle construction where the base or *external* space is instead all of our 3+1D universe/spacetime $\mathcal{M}^{3,1}$, and the total or *internal* 6D space is M^6. The internal space is compactified to obtain de Sitter geometry.

Now, a few conditions must be met to study *spinors* (elements of complex vector spaces that describe particles of half-integer spin) over an expanding spacetime bundle. An associated complex vector bundle, known as a *spinor bundle*, is required. We define spinors using the spin representation of the Lorentz group, as *sections* of the spinor bundle. Sections are the union of different fibre pieces at different total space points. Being that spacetime is a curved space, a connection is required to study how the derivatives of fields transform in different parts of the space or fibre. This is given by the *covariant derivative*, which acts on these spinors under the following representation:

$$D_\mu = \partial_\mu + \frac{1}{4}\omega_\mu^{ab}\gamma_{ab}. \tag{1}$$

Here, ∂_μ is the usual spacetime derivative, ω_μ^{ab} is the *spin connection*, $\gamma_{ab} = \gamma_{[a}\gamma_{b]}$ is the anti-symmetrization of the γ-matrices, and together this represents how the spinor field changes on the space. Similarly to selecting a locally flat patch in a curved manifold, it is useful to work in a locally flat inertial frame of the curved spacetime. This frame is provided by the *frame bundle*: a principal bundle (a fibre bundle possessing group action on a fibre space) that is associated with our spinor bundle. With the frame bundle, we can attach a local frame or coordinate basis to each fibre in the spinor bundle. This allows us to describe the space locally with flat, expanding coordinates given by $tetrad\ coordinates\ e_a^\mu$. These coordinates are sections of the frame bundle that have a Lorentz index μ and a basis index a. Through tetrads, we can write our metric as locally related to the flat Minkowksi metric, bypassing the need for local trivializations (to work in flat space). We can notably define γ-matrices in curved spacetime as:

$$\Gamma^\mu(x) = e_a^\mu(x)\gamma^a, \tag{2}$$

where γ^a are the spatially-constant γ-matrices and e_a^μ is the local tetrad coordinate basis. This provides a notion of spatially-dependent γ-matrices on curved spacetime. The dynamics of the spinors ψ in a curved, expanding spacetime, with the γ-matrices and covariant derivatives, are thus given by the following action:

$$S = \int d^4x\sqrt{-g}\left[\bar{\psi}\Gamma^\mu D_\mu\psi - m\bar{\psi}\psi\right]. \tag{3}$$

Here, $\sqrt{-g}$ is the root determinant of the metric (the negative sign yields a real determinant), which keeps the measure Lorentz-invariant, ψ are the spinor fields, $\bar{\psi} = \psi^\dagger\gamma^0$ are conjugate spinor fields, Γ^μ are the γ-matrices in curved spacetime, D_μ is the covariant derivative associated with the spin connection ω_μ^{ab}, and m is the mass of the spinors. It is sufficient to show that we cannot describe spinors on a curved, expanding spacetime by studying how they fail in a flat, expanding spacetime. To understand the dynamics of these spinor fields, we observe how the expanding metric contributes to the action S on this flat space. Consider the *Friedmann–Lemaître–Robertson–Walker (FLRW)* metric, the simplest metric for a flat expanding universe:

$$ds^2 = -dt^2 + a^2(t)\delta_{ij}dx^i \otimes dx^j. \tag{4}$$

In this equation, $a(t)$ is the scale factor that determines how the universe expands, δ_{ij} is the flat metric for the spatial component of the space, $\{dx^i\}$ is the differential form basis of the cotangent space, and $\otimes$ is the tensor product between forms. To simplify calculations, one could relate this to the Minkowski metric under a conformal transformation. Replacing time t with conformal time η (through $dt = a(\eta)d\eta$) results in the following metric:

$$ds^2 = a^2(\eta)(-d\eta^2 + \delta_{ij}dx^i \otimes dx^j). \tag{5}$$

We can compactly write this conformal relation between the FLRW metric $g^{\mu\nu}$ and the Minkowski metric $\eta^{\mu\nu}$ as $g^{\mu\nu} = a^{-2}\eta^{\mu\nu}$. Given that Minkowski space is flat, the spin connection ω_μ^{ab} vanishes, reducing the covariant derivatives D_μ to normal spacetime derivatives ∂_μ. The covariant derivatives therefore act on the spinor fields as $D_0\psi = \partial_0\psi \equiv \dot{\psi}$ and $D_i\psi = \partial_i\psi \equiv \psi_i'$. Here, ψ_i' is written to mimic usual notation for spatial derivatives such as $f'(x) = \partial_x f$, and the index is not to be confused with the ψ_ν vector-spinors. Accordingly, the action becomes:

$$S = \int d^4x\, a^2\left[\bar{\psi}\left(\Gamma^i\psi_i' - \Gamma^0\dot{\psi}\right) + \frac{1}{4}\bar{\psi}\left(\Gamma^i\omega_i^{ab} - \Gamma^0\omega_0^{ab}\right)\gamma_{ab}\psi - ma^2\bar{\psi}\psi\right], \tag{6}$$

where the Dirac conjugate is used for curved spacetime $\bar{\psi} = \psi^\dagger\Gamma^0$. Although we are in flat space, we keep the vanishing spin connection terms to emphasize the structure of the action when we split spatial and temporal indices. Now, to conceive of spinors on an expanding spacetime, we introduce an *auxiliary field* $\chi = a(\eta)\psi$. Taking the spatial and temporal derivatives of χ and relating them to the original spinor results in the following relations:

$$\dot{\psi} = \frac{\dot{\chi}}{a} - \frac{\dot{a}}{a^2}\chi, \quad \psi_i' = \frac{\chi_i'}{a} - \frac{a_i'}{a^2}\chi. \tag{7}$$

Plugging the expressions for $(\psi, \dot{\psi}, \psi_i')$ in the action, and following algebraic manipulations, we obtain the following:

$$S = \int d^4x\left[\bar{\chi}\left(\Gamma^i\chi_i' - \Gamma^0\dot{\chi}\right) + \frac{1}{4}\left(\Gamma^i\omega_i^{ab} - \Gamma^0\omega_0^{ab}\right)\gamma_{ab}\chi - ma^2\bar{\chi}\chi + \frac{1}{a}\bar{\chi}\left(a_i'\Gamma^i + \dot{a}\Gamma^0\right)\chi\right]. \tag{8}$$

When comparing Equations 6 and 8, we see that the action has developed an extra piece in the form $\frac{1}{a}\bar{\chi}\left(a_i'\Gamma^i + \dot{a}\Gamma^0\right)\chi$. Note that, since $a = a(\eta)$ is not a function of the spatial component of the spacetime, $a_i' = 0$. The extra piece is then just $\frac{\dot{a}}{a}\bar{\chi}\Gamma^0\chi$, and expanding $\Gamma^0 = e_{0a}\gamma^a$ yields $\frac{\dot{a}}{a}\bar{\chi}e_{0a}\gamma^a\chi$. This extra term modifies the equations of motion of the auxiliary spinors. Varying the action functional with respect to the fields $(\chi, \bar{\chi}, e_a^\mu)$ gives us the following set of equations of motion (EOM):

$$\frac{\delta S}{\delta\bar{\chi}} = \partial_\mu\left(\bar{\chi}\Gamma^\mu\right) + \bar{\chi}\left(\frac{\dot{a}}{a}\Gamma^0 - ma^2\right) = 0, \tag{9}$$

$$\frac{\delta S}{\delta\bar{\chi}} = \Gamma^\mu\partial_\mu\chi + \left(\frac{\dot{a}}{a}\Gamma^0 - ma^2\right)\chi = 0, \tag{10}$$

$$\frac{\delta S}{\delta e_a^\mu} = \bar{\chi}\gamma^a\left(\frac{1}{a}\partial_\mu a + \partial_\mu\chi\right) = 0. \tag{11}$$

When integrating by parts, the variations of the form $\delta(\partial_\mu\chi) = \partial_\mu(\delta\chi)$[12] allow us to disregard the boundary terms as the FLRW universe is asymptotically flat, and the fields $(\chi, \bar{\chi})$ vanish at infinity. In the equations of motion for $(\chi, \bar{\chi})$, the extra terms proportional to $\frac{\dot{a}}{a}\Gamma^0$ cause the Fourier expansions of the EOM solutions to develop time-dependent frequencies $\omega(t)$. However, these time-dependent frequencies pose a challenge as they prevent the definition of a vacuum. Similarly to scalar fields, the expansion of χ into Fourier modes includes creation and annihilation operators $(a_k^\dagger, a_k)$, which are dependent on the frequency of the spinor. This poses a problem because the time-dependent annihilation operators, which annihilate the vacuum at some point in time, do not necessarily annihilate the vacuum at another time. Without a well-defined notion of a vacuum, describing fields or particles in the expanding spacetime becomes impossible. Furthermore, defining an action for a quantum field theory at a given energy scale requires an *effective Wilsonian action*. This is an action whereby

irrelevant operators have been removed after integrating out high energy modes. However, if the fields develop time-dependent frequencies, one cannot define an effective Wilsonian action by integrating out high energy spinor modes to keep observables finite. This is because low energy modes can evolve in time to become high energy modes. This is why we cannot define vacuum states over an expanding curved geometry such as de Sitter space.

More problems occur when considering a UV-complete theory that has supersymmetry. It has been shown that over an expanding spacetime, bosonic fields develop time-dependent masses[10]. However, in our case of fermionic fields, their mass remains constant while their frequencies become time-dependent. However, in the supersymmetric framework, the zero-point energy no longer cancels out. This means that we would have to perturb the system about a diverging vacuum energy, meaning we have no reference of finite observable energy. Furthermore, this explicitly breaks supersymmetry, so the theory does not agree with our compactifications coming from type IIB string theory. To prevent these problems, we make use of excited coherent states over supersymmetric Minkowski space, which only break supersymmetry spontaneously, meaning that the vacuum remains supersymmetric.

Excited Coherent States

Here, we alleviate problems associated with fields defined over expanding spacetimes in a UV-complete theory by considering coherent states over supersymmetric Minkowski space. In this framework, we will show that de Sitter space can be a result of compactifications of type IIB string theory and that one calculates a positive cosmological constant for an expanding universe. These coherent states, known as *Glauber-Sudarshan* (GS) states $|\sigma\rangle$, expand and share the isometries of de Sitter space. As these states share the same isometries as the de Sitter space metric, they are of particular interest to replace the problematic interacting de Sitter vacuum states (or *vacua*) $|\Omega\rangle$. The idea is to use these states to compute quantum corrections to the metric *operator* $\hat{g}_{\mu\nu}$, where part of its representation contains an expression for the cosmological constant, which will turn out positive.

We must first consider why coherent states are used in quantum field theory (QFT) and elaborate on the nature of Glauber-Sudarshan states. Additionally, we should understand the nature of the supersymmetric Minkowski space that we are working over. Then, we will move onto considering the metric quantum corrections.

Glauber-Sudarshan (GS) States

For a *free QFT* with a single bosonic Degree of Freedom (DOF) ($\alpha_1 \equiv \alpha$ for example), the coherent state $|\alpha\rangle$ is a shift of the free vacuum $|\alpha\rangle = \mathbb{D}_0(\alpha)\,|0\rangle = \exp\left(\alpha a_k^\dagger - \alpha^* a_k\right)|0\rangle$, where $\mathbb{D}_0(\alpha)$ is the unitary displacement operator, α is a complex number, and $(a_k^\dagger, a_k)$ are the usual creation and annihilation operators. These states are useful because they preserve the degree to which a quantum system exhibits wave-like behavior, such as interference and diffraction. Furthermore, they are the excited quantum states that most closely resemble classic states, with minimal uncertainty in position and momentum. This is useful to us as our theory must reproduce classical physics in certain limits. In the case for an *interacting QFT*, however, with multiple bosonic DOFs $\{\alpha_i\} \equiv \sigma$, we instead have $|\sigma\rangle = \mathbb{D}(\sigma,t)\,|\Omega\rangle$ — the Glauber-Sudarshan states. Being that the "vacuum" state we utilize is now *excited*, there is an ambiguity of the displacement operator and it becomes non-unitary. For the case of interacting vacua, the structure is not as trivial as the free vacua, so preserving the unitarity of the shift operator is difficult. It is however still possible to represent it in terms of the interacting Hamiltonian over a temporal domain as follows[13]:

$$\mathbb{D}(\sigma,t) = \lim_{T\to\infty(1-i\epsilon)} \mathbb{D}_0(\sigma,t)\exp\left(iM_p\int_{-T}^{t} dt\ \mathbf{H}_{\text{int}}\right). \tag{12}$$

Here, $\mathbb{D}_0(\sigma,t)$ is a time-dependent unitary displacement operator for the free vacuum, $\mathbf{H}_{\text{int}}$ is the interaction Hamiltonian of the theory, M_p is the Planck mass, and the limit is slightly in the imaginary direction $(1-i\epsilon)$ for the same reason as it is for propagators: to avoid poles in the complex phase that give rise to divergences in observables. Now, for reasons which will become clear later, we must *uplift* type IIB string theory to *M-theory*, an 11D string theory (we can recover type IIB string theory by compactifying along the 11th dimension in M-theory). The multiplet for the fields in M-theory is given by $(g_{ab}, C_{abc}, \psi_a)$, where g_{ab} is the metric field, C_{abc} is 3-form field, and ψ_a is a vector-spinor. One can imagine describing all the DOFs of the respective fields via a collection of numbers (in the form of a vector / matrix / higher-order construction). The DOFs of the metric field g_{ab} are captured by the set $\{\alpha_{ab}\}$ for some numbers α_{ab}, the set $\{\beta_{abc}\}$ for some numbers β_{abc} captures the DOF of the 3-form field C_{abc}, and finally the set $\{\gamma_a\}$ for some numbers γ_a captures the DOF of the vector-spinor field ψ_a. Thus, fields in the M-theory multiplet respectively come with collective DOFs $(\{\alpha_{ab}\},\{\beta_{abc}\},\{\gamma_a\})$ (128 from the bosonic sector and 128 from the fermionic sector, totaling 256 DOFs). Denoting all the DOFs as $\sigma \equiv (\{\alpha_{ab}\},\{\beta_{abc}\},\{\gamma_a\})$, we write the coherent Glauber-Sudarshan state as[14]:

$$|\sigma\rangle \equiv \mathbb{D}(\sigma,t)\,|\Omega\rangle = \bigotimes_k \left(\sum_{f_k} \Psi^\sigma(f_k)\,|f_k\rangle\right). \tag{13}$$

Here, the sum is over all the fields $f_k = (\{g_{ab}(k)\},\{C_{abc}(k)\},\{\psi_a(k)\})$ and the tensor product is over the mode momenta k. Ψ^σ is the wave function of the GS state (corresponding to the DOFs of the field f_k given by σ) which can be seen as a product of normalized Dirac delta functions, and $|f_k\rangle$ are eigenstates of the momentum wavefunction of the GS states in the configuration space. This is a general state that contains all 256 DOFs of the field in M-theory and is coherent. This is the description of the states we will use for the quantum corrections.

Supersymmetric Minkowski Space

Since type IIB theory is a 9+1D theory, we must compactify on a 6D manifold to recover our 3+1D universe (the de Sitter vacuum solution). For explicit calculations, it is useful to pick a *slicing* of de Sitter space (a selection of a foliation). We choose one which makes de Sitter space appear flat within a certain region, given by $ds^2 = 1/(\Lambda|t|^2)\eta_{\mu\nu}dx^\mu \otimes dx^\nu$. Here, t is the conformal time coordinate (instead of writing η), Λ is the cosmological constant, $\eta_{\mu\nu}$ is the usual Minkowski metric, and dx^μ is the basis of the contangent space. A flat slicing gives us a temporal domain $-1/\sqrt{\Lambda} \le t < 0$, over which the metric is well defined, where $t = 0$ represents late times. This temporal domain comes precisely from the trans-Planckian censorship conjecture (TCC) time scale[15]. Although we are working with supersymmetric Minkowski space (a superspace), it suffices for our analyses to look at the non-supersymmetric part of the 10D space. In this case, the full space is given by:

$$\mathcal{M}_{10} = \mathbb{R}^{3,1} \times \mathcal{M}_4 \times \mathcal{M}_2, \tag{14}$$

where $\mathbb{R}^{3,1}$ is 3+1D Minkowski space, and $\mathcal{M}_4 \times \mathcal{M}_2 \equiv \mathcal{M}_6$ is some non-Kähler 6D internal space (written as a product space to account for multiple scaling factors for different pieces of the internal space). Moreover, to align

with more general geometries that include warping, we define the metric of the 10D space as the warped geometry[13]:

$$ds^2 = \frac{1}{\Lambda H^2(y)|t|^2}\eta_{\mu\nu}dx^\mu \otimes dx^\nu + H^2(y)\left[F_1(t)g_{\alpha\beta}dy^\alpha \otimes dy^\beta + F_2(t)g_{mn}dy^m \otimes dy^n\right]. \quad (15)$$

Here, $H(y)$ is the *warp factor* that depends on the internal space coordinates $\{y^m, y^\alpha\}$, and which makes the external space explicitly dependent on the internal space's behaviour. $\{x^\mu\}$ are the coordinates of the external space $\mathbb{R}^{3,1}$. This makes us no longer consider a local quantum field theory as it includes non-local interactions. Furthermore, $(F_1(t), F_2(t))$ are scaling factors of the different internal subspaces $(\mathcal{M}_4, \mathcal{M}_2)$, respectively, equipped with metrics $(g_{\alpha\beta}, g_{mn})$. When computing quantum corrections to the metric operator, we compute path integrals over the flat metric to reproduce the above metric. Path integrals require a full description of the system's action, however, there are currently *no* well-defined actions for type IIB string theory[16]. To work with a string theory that has a well-defined action, we consider the uplift of type IIB string theory to M-theory, an 11D string theory. As mentioned in the previous section, the multiplet for M-theory is $(g_{\mu\nu}, C_{\mu\nu\rho}, \psi_\mu)$, and the 10D space now becomes an 11D space given by the following metric:

$$\mathcal{M}_{11} = \mathbb{R}^{3,1} \times \mathcal{M}_4 \times \mathcal{M}_2 \times \frac{\mathbb{T}^2}{\mathbb{Z}_2}. \quad (16)$$

Notice that the extra 11th dimension appears as the quotient space between the torus $\mathbb{T}^2$ and the $\mathbb{Z}_2$ (the group of integers mod 2). We define coordinates of the space as $(x^\mu, x^\nu) = (x^0, \ldots, x^3)$, $(y^m, y^n) = (x^4, \ldots, x^9)$, and $(\omega^a, \omega^b) = (x^{10})$, for the spaces $(\mathbb{R}^{3,1}, \mathcal{M}_6, \frac{\mathbb{T}^2}{\mathbb{Z}_2})$ respectively. This modifies the overall metric of the 11D space, the explicit form for which is shown in Alexander *et al.*[16] For clarity, we also consider indices $(M, N) = (0, \ldots, 10)$, which takes into account the information (coordinates) of the entire space and will work with the multiplet $(g_{MN}, C_{MNP}, \psi_M)$ defined over all of $\mathcal{M}_{11}$. For explicit computations with this group of fields, the overall multiplet can be decomposed over the different sub-pieces of the spacetime as direct sums of singlets (this is known as a *dimensional reduction*).

Quantum Corrections

Now that we have an expression for the space we are working with and have defined the GS states, we move onto computing the quantum corrections in the form of contributions affecting the metric *operator* $\hat{g}_{\mu\nu}$. To note, the explicit reference to *operators* addresses the subtlety between operators and fields when making use of path integrals. The correction takes form of the expectation value $\langle\sigma|\,\hat{g}_{\mu\nu}\,|\sigma\rangle$ normalized by $\langle\sigma|\sigma\rangle$ (defined as $\langle\hat{g}_{\mu\nu}\rangle_\sigma$). The expression for the correction is a quotient of path integrals (functional integrals over all possible evolutions of the fields) over the M-theory multiplet in 11D (ref. 16) (denoted as $\langle\hat{g}_{\mu\nu}\rangle_\sigma$):

$$\frac{\int \mathcal{D}[g_{MN}]\mathcal{D}[C_{MNP}]\mathcal{D}[\psi_M]\mathcal{D}[\bar{\psi}_N]\; e^{iS}\; \mathbb{D}^\dagger(\sigma,t)g_{\mu\nu}(x)\mathbb{D}(\sigma,t)}{\int \mathcal{D}[g_{MN}]\mathcal{D}[C_{MNP}]\mathcal{D}[\psi_M]\mathcal{D}[\bar{\psi}_N]\; e^{iS}\; \mathbb{D}^\dagger(\sigma,t)\mathbb{D}(\sigma,t)}. \quad (17)$$

Here, $\mathcal{D}[A_{M\ldots N}] \sim \prod_{M,\ldots,N} dA_{M\ldots N}$ are the path integral field measures (where $A_{M\ldots N}$ represents the different fields of the multiplet that are integrated over), S is the total action of the system, $\mathbb{D}^\dagger(\sigma,t)$ is the non-unitary shift operator, and $g_{\mu\nu}(x)$ is the metric of Minkowski space $\mathbb{R}^{3,1}$. This path integral is much too complicated at this level of generality with Grassmanian integrals (coming from integrals over the vector-spinors $\mathcal{D}[\psi_M]\mathcal{D}[\bar{\psi}_N]$), so we picked three representative sample scalars $(\varphi_1, \varphi_2, \varphi_3)$ for each of the DOFs of the fields in the M-theory multiplet. Theses scalars fix the DOFs to $\sigma \equiv (\varphi_1, \varphi_2, \varphi_3)$, and so for the DOFs φ_1 that contribute towards the metric operator correction, we have the replacement $\langle\hat{g}_{\mu\nu}\rangle_\sigma \to \langle\varphi_1\rangle_\sigma$. Although one might assume the expectation value of a scalar to vanish, in this case we are still working with the different fields in the M-theory multiplet, so have simply *reduced* their DOFs to that of scalars. Once fixed, the path integrals in the numerator and denominator of $\langle\varphi_1\rangle_\sigma$ fall under a class of path integrals that can be computed using *nodal diagrams*, due to the shifted-vacuum structure of the GS state. The shifted-vacuum structure is in reference to the GS state being a non-trivial shift of the vacuum state $|0\rangle$ to the GS state $|\sigma\rangle$ via the non-unitary shift operator $\mathbb{D}(\sigma,t)$. This shift structure allows us to make use of nodal diagrams, which is the set of Feynman diagrams that capture the information of higher order point functions[14]. These diagrams emerge from how the different momenta of the fields within the multiplet are summed over. Using the amplitudes of the nodal diagrams $\{\mathcal{A}_s\}$, it is possible to express the quantum correction as the following[14] (denoted as $\langle\varphi_1\rangle_\sigma$):

$$\frac{\text{TLN} + \sum_{n,\ldots,s} c_{mnpqrs}\,\mathcal{N}^{(1)}_{nmp}(k;q)\otimes\mathcal{N}^{(2)}_{nmp}(l;r)\otimes\mathcal{N}^{(3)}_{nmp}(f;s)}{\text{TLD} + \sum_{n,\ldots,s} c_{mnpqrs}\,\mathcal{N}^{(1)'}_{nmp}(k;q)\otimes\mathcal{N}^{(2)}_{nmp}(l;r)\otimes\mathcal{N}^{(3)}_{nmp}(f;s)}. \quad (18)$$

Here, (TLN, TLD) are the tree-level Feynman diagram contributions to the numerator and denominator path integrals, respectively. The contributions of Feynman diagrams with loops comes from summing over the amplitudes of the nodal diagrams. Here, $\mathcal{N}^{(i)}_{nmp}(a;b)$ represents the Feynman diagram amplitude (nodal diagram amplitude) of interactions between fields $\varphi^b_i(a)$, where i labels the different scalars that represent the DOFs of the multiplet in M-theory $(\varphi_1, \varphi_2, \varphi_3)$, a labels the incoming momenta of the fields, and b is an integer. The total loop contribution comes from summing over the nodal amplitudes for different i, which is weighted by the coupling constants c_{nmpqrs}. Furthermore, the prime in $\mathcal{N}^{(1)'}_{nmp}(k;q)$ refers to interaction diagrams without field sources. These nodal diagrams diverge with structure of the Gevrey kind (meaning they diverge factorially) and require Borel resummation[‡] to restructure the divergence into non-perturbative solitonic corrections. Using the nodal amplitudes and Borel re-summation gives the following correction to the metric operator[16]:

$$\langle\varphi_1\rangle_\sigma = \sum_{\{s\}}\left[\frac{1}{g^{1/l}_{(s)}}\int_0^\infty d\mathcal{B}\exp\left(-\frac{\mathcal{B}}{g^{1/l}_{(s)}}\right)\frac{1}{1-\mathcal{A}_{(s)}\mathcal{B}^l}\right]_{P.V} \times \int_{k_{\text{IR}}}^{\mu} d^{11}k\;\frac{\bar{\alpha}_{\mu\nu}(k)}{a(k)}\;\mathbf{Re}\left(\psi_k e^{-i(k_0-\bar{\kappa}_{\text{IR}})t}\right), \quad (19)$$

where $\{s\}$ is the set of interactions, $g_{(s)}$ is the set of coupling constants, $\mathcal{B}$ parametrizes an axis in the Borel plane, $\mathcal{A}_{(s)}$ is the amplitude of all possible nodal diagrams, $P.V$ is the principal value of the integral over $\mathcal{B}$, $(k_{\text{IR}}, \bar{\kappa}_{\text{IR}})$ are IR scales, $\bar{\alpha}_{\mu\nu} = \alpha_{\mu\nu}/V$ is a 2-form normalized by the volume V of the space $\mathcal{M}_{11}$, $a(k) = k^2/V$, l is the total amount of fields minus one, k is the momentum (which has an associated k_0 component), and ψ_k is the spatial wavefunction of the GS state over the space $\mathcal{M}_{11}$ (projecting the GS state $|\sigma\rangle$ into the coordinate space of $\mathcal{M}_{11}$). As before, $(\mu,\nu) = (0,\ldots,3)$

[‡] A resummation is a prescription to make a specific class of divergent series convergent via regularization and rescaling.

labels the spacetime coordinates of the external space $\mathbb{R}^{3,1}$. It turns out that the first piece of the form $\sum[\dots]$ exactly corresponds to the inverse of the cosmological constant (at a given IR scale κ)[14], which is defined over the same time domain as the GS state ($-\frac{1}{\sqrt{\Lambda}} < t \leq 0$). We then have:

$$\frac{1}{\Lambda^{\kappa}} \equiv \sum_{\{s\}} \left[\frac{1}{g_{(s)}^{1/l}} \int_0^{\infty} d\mathcal{B} \exp\left(-\frac{\mathcal{B}}{g_{(s)}^{1/l}} \right) \frac{1}{1 - \mathcal{A}_{(s)}\mathcal{B}^{l}} \right]_{P.V}, \quad -\frac{1}{\sqrt{\Lambda}} < t \leq 0. \tag{20}$$

This is known as the *integral form* of the cosmological constant Λ^{κ}, and its integral in this case is **positive definite** over the flat slicing of the temporal domain[16]. This means that the constant affords a positive sign — the desired sign of the cosmological constant. To recover the result for type IIB string theory from M-theory, one need only compactify along a compact direction, which leaves the above form of the cosmological constant unchanged. This shows that using the GS states (which have the isometries of Minkowski space and preserve supersymmetry) yields the correct sign for the cosmological constant. Here, the cosmological constant functional expression is issued from the quantum corrections of the metric operator, and its correct positive sign indicates that we indeed have compactifications of the de Sitter vacuum solution from type IIB string theory. Hence, with the combination of GS states and Minkowski space, we can consider de Sitter space as a candidate of our universe in a UV-complete theory (as coming from a type IIB string theory compactifications).

Conclusion

The swampland scenario prevents expanding spaces with positive curvature from producing a cosmological constant with the correct sign (i.e., positive), whether in UV-complete or incomplete theories. Furthermore, in the presence of an expanding spacetime, fields (and hence the particles they describe) develop time-dependent frequencies, meaning particles are not well defined.

To counteract this, we considered a candidate for our universe, de Sitter space, in a UV-complete theory whereby we recovered the de Sitter vacuum solution through compactifications of type IIB string theory. To avoid the ambiguity of de Sitter vacuum states over an expanding spacetime, we reviewed work done on a class of general coherent states (known as Glauber-Sudarshan states, which have the desired de Sitter isometries) over supersymmetric Minkowski space. We showed that we can define these states over a given temporal domain via the non-unitary shift operator which excites the interacting vacua to give the GS states. These states allowed us to compute the metric operator quantum corrections, which contains an expression for the cosmological constant. We showed that the constant is positive definite, meaning it has the correct sign for an expanding universe. This shows that we can have consistent compactifications to de Sitter space in type IIB string theory, and thus the de Sitter vacuum solution can be used as a candidate for our universe.

Acknowledgements

I would like to thank Simon Caron-Huot, Brent Pym, Evan McDonough, Mathieu Boisvert, Clément Virally, Guillaume Payeur, Bohdan Kulinich, and David leNir for our discussions related to string theory, algebraic topology, and string cosmology. I would also like to thank Jonathan Eid, Alexander Kroitor, and Hans Hopkins for reading and giving feedback on previous drafts of this paper. Additionally I would like to thank Keshav Dasgupta for (among many other things) teaching a superstring theory course over 9 months which helped immensely in catching up to speed and understanding the framework of string theory, prior to which (and hence prior to the thesis) I knew nothing of.

References

1. Cohen, A. G., Kaplan, D. B. & Nelson, A. E. Effective field theory, black holes, and the cosmological constant. *Phys. Rev. Lett.* **82,** 4971–4974 (1999). https://doi.org/10.1103/PhysRevLett.82.4971
2. Munkres, J. R. *Topology* (Pearson, 2014).
3. Kachru, S., Kallosh, R., Linde, A. D. & Trivedi, S. P. De Sitter vacua in string theory. *Phys. Rev. D* **68,** 046005 (2003). https://doi.org/10.1103/PhysRevD.68.046005
4. Douglas, M. R. Dirichlet branes, homological mirror symmetry, and stability. Preprint at https://doi.org/10.48550/arXiv.math/0207021 (2002).
5. Danielsson, U. H. & Van Riet, T. What if string theory has no de Sitter vacua? *Int. J. Mod. Phys. D* **27,** 1830007 (2018). https://doi.org/10.1142/S0218271818300070
6. Cicoli, M. et al. String cosmology: From the early universe to today. *Phys. Rept.* **1059,** 1–155 (2024). https://doi.org/10.1016/j.physrep.2024.01.002
7. Anderson, P. R., Mottola, E. & Sanders, D. H. Decay of the de Sitter Vacuum. *Phys. Rev. D* **97,** 065016 (2018). https://doi.org/10.1103/PhysRevD.97.065016
8. Maldacena, J. M. & Nunez, C. Supergravity description of field theories on curved manifolds and a no go theorem. *Int. J. Mod. Phys. A* **16,** 822–855 (2001). https://doi.org/10.1142/S0217751X01003937
9. Dasgupta, K., Gwyn, R., McDonough, E., Mia, M. & Tatar, R. de Sitter Vacua in Type IIB String Theory: Classical Solutions and Quantum Corrections. *J. High Energy Phys.* **07,** 054 (2014). https://doi.org/10.1007/JHEP07(2014)054
10. Mukhanov, V. & Winitzki, S. *Introduction to Quantum Effects in Gravity* (Cambridge University Press, 2007). https://doi.org/10.1017/cbo9780511809149
11. Baez, J. & Muniain, J. P. *Gauge Fields, Knots and Gravity* (World Scientific, 1994). https://doi.org/10.1142/2324
12. Carroll, S. M. *Spacetime and Geometry* (Cambridge University Press, 2019). https://doi.org/10.1017/9781108770385
13. Brahma, S., Dasgupta, K. & Tatar, R. Four-dimensional de Sitter space is a Glauber-Sudarshan state in string theory. *J. High Energy Phys.* **2021** (2021). https://doi.org/10.1007/jhep07(2021)114
14. Brahma, S. et al. Resurgence of a de Sitter Glauber-Sudarshan State: Nodal Diagrams and Borel Resummation. *Fortsch. Phys.* **2023,** 2300136 (2022). https://doi.org/10.1002/prop.202300136
15. Bedroya, A. & Vafa, C. Trans-Planckian Censorship and the Swampland. *J. High Energy Phys.* **09,** 123 (2020). https://doi.org/10.1007/JHEP09(2020)123
16. Alexander, S., Dasgupta, K., Maji, A., Ramadevi, P. & Tatar, R. de Sitter State in Heterotic String Theory. Preprint at https://doi.org/10.1002/prop.202400163 (2023).