VOLUME II (2017)

journals.mcmaster.ca/iScientist

THE iSCIENTIST

McMASTER'S UNDERGRADUATE INTERDISCIPLINARY SCIENCE JOURNAL

THE *i*SCIENTIST
McMASTER'S UNDERGRADUATE INTERDISCIPLINARY SCIENCE JOURNAL
Volume II, 2017

EDITORIAL BOARD

Alexander Dhaliwal & Kyra Simone
SENIOR EDITORS

Laura Green & Varsha Jayasankar
LEVEL III EDITORS

Elsie Loukiantchenko & Sabrina Macklai
LEVEL II EDITORS

Dr. Sarah Symons & Andrew Colgoni
EDITORS-IN-CHIEF

James Lai
SENIOR ADVISOR

Effective communication of research in a variety of contexts is an extremely important aspect of conducting oneself as a professional scientist. *The iScientist* is an interdisciplinary journal that provides the unique opportunity to experience the scientific publication process as an undergraduate student in a well-organized and professional setting. By submitting to *The iScientist*, students can develop a familiarity with the publication process through reporting their original research, accessing and discussing scientific literature, participating in the peer review process, and writing their own academic research papers.

DEAR READERS

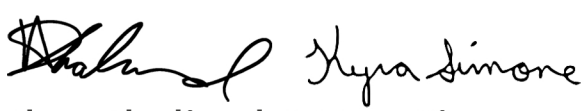
Thank you for taking the time to pick up, flip through, and curiously ponder the letters and articles of *The iScientist* Vol. II. This year, we are delighted to showcase topics ranging from mathematical models of opinion dynamics to viral transneural tracing to Galilean moon exploration. The articles that are included in this journal represent students from a diverse range of backgrounds and fields of science, embarking upon research across a breadth of disciplines. Whether they are extending the frontiers of their chosen scientific concentration or exploring new areas of interest, this journal encourages science students to discover more about their world and produce high-quality, accessible pieces of writing to share with their community.

At its core, *The iScientist* serves as an opportunity for students to experience a professional-level journal experience at an undergraduate level. Since the journal's inception, we have fought to maintain a high standard for the pieces of work that we feature. This involves multiple rounds of peer review to ensure that work is meticulously checked and referenced, fostering a community of academic integrity and robustness. Although it would be easier to loosen these stringencies to allow a wider variety of articles into our issues, we feel a responsibility to authors and readers to feature only top-quality undergraduate work, refined by a rigorous peer review process.

This year brings a wealth of new changes to our journal. Perhaps most substantially, our Editorial Board was fortunate enough to receive funding to publish physical volumes for the next three years, a result only attainable through the generous support of the McMaster Science Society's Science Initiative Fund. In addition to having paper copies of the journal, we have also begun to lay the foundations for a series of workshops that emphasize scientific literacy and publishing at the undergraduate level. These workshops will begin in the 2017/2018 year, and we hope for them to become a pillar function of our board. In this vein, we hope to continue to build a community here at McMaster where students are encouraged to explore interdisciplinary science and display their findings in a professional, worthwhile manner.

We hope that you continue to demand the most from this Editorial Board. In return, we will continue to work hard to showcase the most inspiring, thought-provoking, and original research that students of this School can produce. The success of our work truly is in your hands, and so we encourage you to continue what you do best: scoff at siloed scientific intentions and embrace a scientific paradigm without borders.

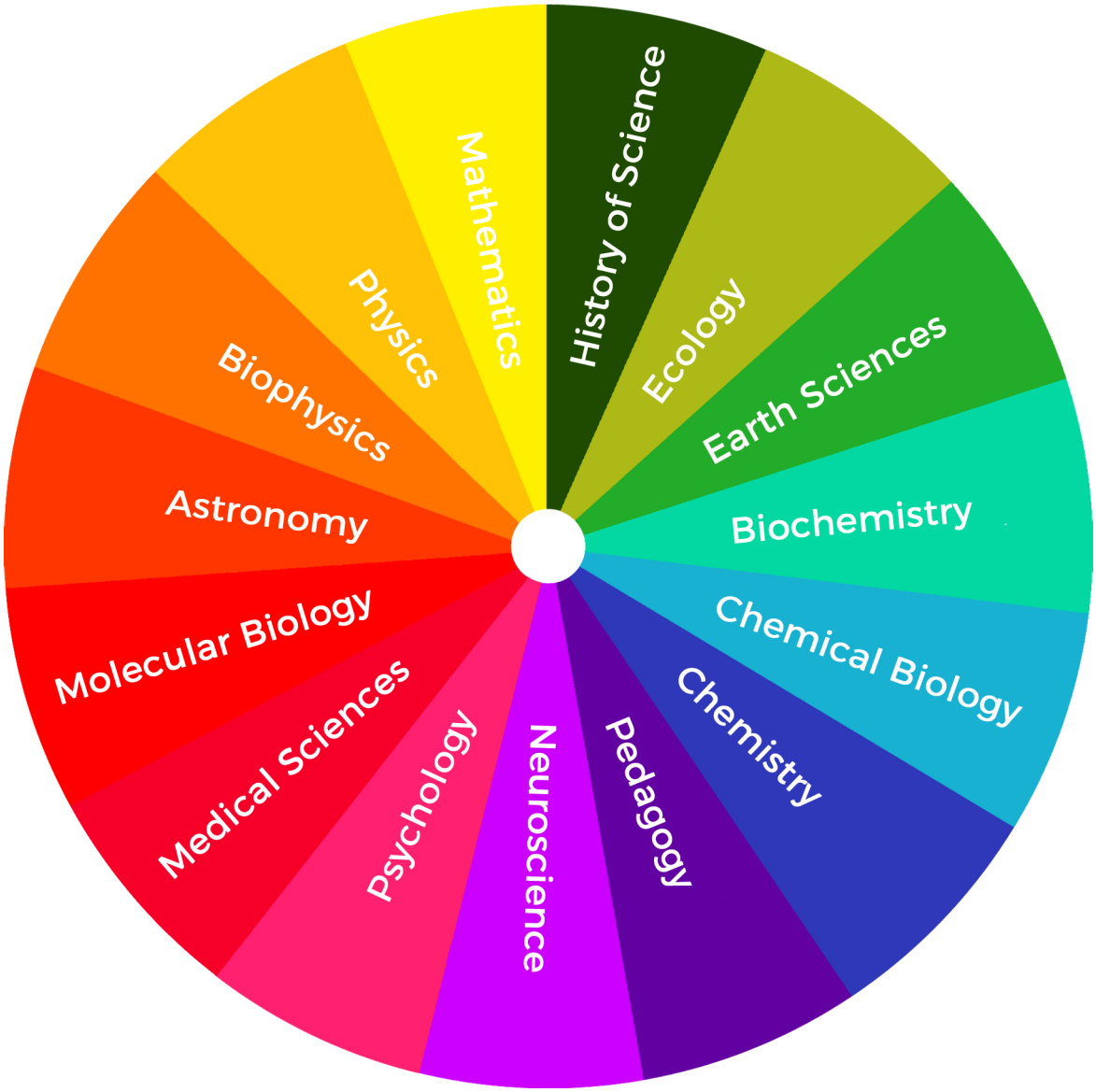
All the best,



Alex Dhaliwal & Kyra Simone
Senior Editors of *The iScientist*

Reading THE *i*SCIENTIST

The iScientist is an interdisciplinary journal. Our articles are divided into fifteen disciplines, colour-coded using this scheme.



The iScientist publishes five types of articles

Original Research
Articles centred around independent research conducted by the authors

Literature Review
A detailed review of previous research on a specific topic

Mathematical Paper
Studies regarding theoretical considerations or applications of mathematical concepts

History of Science
A historical perspective of a topic within science or technology

Letter to *The iScientist*
Short pieces that discuss scientific research or argue an opinion

TABLE OF CONTENTS

1 **The Induction of the Jasmonic Acid (JA) Plant Defence Pathway Affecting the Host-Plant Selection of *Myzus persicae***
Rama Al-Atout, Chimira Andres, Michael Chong, Ross Edwards, Julia Pantaleo

7 **Comparison of Mathematical Models of Opinion Dynamics**
Aurora Basinski-Ferris

15 **Secrets in Math: How do we Fly?**
Simon Zhang

22 **Emerald Ash Borer Incidence and Infestation at McMaster Forest Teaching and Research Facility**
Kyra Simone

28 **Current Understanding of Europa and Potential in Upcoming Exploration**
James Lai

31 **Using Viruses as Molecular Biology Tools: A Review of Viral Transneural Tracing**
Jacqueline Watt

41 **The Effects of Salinity and Activated Charcoal on the Herbivory of *Arabidopsis thaliana* by *Myzus persicae***
Sommer Chou, Alexi Doan, Helena Koniar, Bryce Norman

47 **Investigating the Relationship between Variance of Transition Temperatures**
Angelico Obille

52 **Human Papilloma Virus: Combatting Cancer with Vaccination**
Dakota Binkley, Jacqueline Watt, Achint Bajpai

The Induction of the Jasmonic Acid (JA) Plant Defence Pathway Affecting the Host-Plant Selection of *Myzus Persicae*.

RAMA AL-ATOUT, CHIMIRA ANDRES, MICHAEL CHONG, ROSS EDWARDS, JULIA PANTALEO
Integrated Science Program, Class of 2018, McMaster University

SUMMARY

Host selection in aphids such as *Myzus persicae* is an intriguing research pathway which investigates aphid population interactions within ecological systems. Induced plant resistance activates an effective signalling pathway, the jasmonic acid pathway, and in turn affects the performance of aphids. In this study, we look to study the short-term host selection and preference of *M. persicae* between wounded and healthy wild-type *Arabidopsis* plants. We conducted a microcosm experiment wherein aphid movement between neighbouring plants was observed over a 12-day period. Each microcosm contained one of four possible combinations of damaged and undamaged host plants, and three *M. persicae*. Host-plant preference was measured using the percentage of *M. persicae* that moved to or from damaged plants and analyzed using pairwise t-tests between treatments. *M. persicae* showed a preference for undamaged hosts over damaged hosts.

Received: 02/27/2016 **Accepted:** 02/20/2017 **Published:** 07/31/2017
URL: https://journals.mcmaster.ca/iScientist/article/view/1152

Keywords: *Arabidopsis thaliana*, *Myzus persicae*, green peach aphid, host preference, metapopulation dynamics, herbivory, mechanical wounding, defence pathways

INTRODUCTION

Host selection in aphids such as *Myzus persicae* is an intriguing research pathway that investigates aphid population interactions within ecological systems. Understanding the underlying mechanism of aphid metapopulation dynamics and how that is affected by the type and condition of the plant is crucial for predicting their movements between local crops or plants in general.

METAPOPULATION DYNAMICS

Measuring aphid population movement is crucial for assessing the impact of generalist species like *M. persicae* on a species of plant, especially when introduced into a new environment. Seasonal migration is understood in the context of gene-environment interactions, in particular, how annual climatic change and day length influences genotype frequencies in different parts of the world. A study by Blackman in 1974, for example, reported significant differences in the sexual reproduction and emigration of *M. persicae* populations that correlated directly to temporal and climatic changes. This research, however, does not consider the immediate "decision-making" of *M. persicae.* Past research has investigated how the species of a host plant can affect host selection, as well as the factors that influence relocating to a new host plant between healthy plants (Sutherland and Mittler, 1971; Renzo, 2005) However, it is unknown whether mechanical damage to the plant plays a role in host selection.

APHID HOST PREFERENCE AND PROBING

Previous research has attempted to identify factors affecting aphid host preference. In an isolated

system, we expect "plant crowding" to be the primary regulating factor and main reason for aphid migration between plants. Plant crowding occurs when multiple individuals in a population begin to compete for resources on the same plant (Sutherland and Mittler, 1971). However, other factors are known to affect the suitability of a host. All aphids exhibit some kind of leaf-probing behaviour before selecting a host plant (Goggin, 2007). When an aphid probes a plant, it searches for cues such as pH, protein content, and sugar content to inform it if the chosen feeding spot is viable (Elzinga, et al., 2013). In addition to phloem sampling cues, present volatiles may also affect host preference (Renzo, et al., 2005). Depending on food source quality and plant defences, an aphid may migrate to a neighbouring plant and restart the probing process.

PLANT DEFENCE

The jasmonic acid (JA) plant defence pathway, which is induced by outer cell tissue damage of a plant, can be harmful to aphids (Jaouannet, et al., 2014). Lipoxygenase, or LOX, is one of the first enzymes involved in the biosynthesis of JA (Moran, 2001). Leaf wounding as well as aphid feeding have been found to induce the expression of LOX genes, suggesting that plant sensitivity to phloem feeding can involve JA synthesis (Hyun et al., 2008), and the build-up of JA and methyl jasmonate, a methyl ester of JA (He, et al., 2002). The formation and accumulation of JA stimulate the induction of genes for proteinase inhibitors, which have been found to deter insect feeding by inhibiting major insect digestive enzymes (Rockwood, 2015), delaying larval development, and negatively affecting adult insect herbivores. This is supported by a study by Thaler, et al. (2001), which found that the presence of JA increased the activity of proteinase inhibitors and polyphenol oxidase, thus causing a consequent decrease in herbivory.

Given the negative effect of the JA defence pathway on aphid populations, it is hypothesized that aphids are less inclined towards mechanically damaged host plants due to the presence of metabolites from the induced JA pathway. In this experimental study, our goal was to understand whether *M. persicae* has a preference between mechanically damaged and undamaged host plants by analyzing the short-term migratory behaviour of *M. persicae* as a result of *A. thaliana*'s induced plant defence response. Specifically, our experiment attempted to determine if the induced JA pathway via mechanical wounding in *Arabidopsis thaliana* causes aphid relocation to an undamaged host plant.

In order to answer these questions, a controlled microcosm experiment using the interacting species *M. persicae* and *A. thaliana* was conducted wherein different combinations of damaged and undamaged plants were used to observe aphid host selection behaviour in each microcosm. This experimental design was used to answer the question: how does the JA pathway induced by mechanical wounding influence the short-range host selection of *M. persicae* using *A. thaliana* as a model host plant?

RESULTS

Since aphid populations varied among microcosms, the primary measure used to determine aphid host preference was the mean percentage of the total aphid population that migrated to Plant 2 in each microcosm – Plant 2 being the plant adjacent to Plant 1 where the aphids were originally placed. Only data from day 12, the final day, were used because analysis of the proportions over time would not provide additional useful information about the research question, and sufficient time was to be provided for *M. persicae* to move to the adjacent plant. Since the damaged-undamaged (Plant 1 was cut, Plant 2 was not cut) and undamaged-damaged (Plant 1 was not cut, Plant 2 was cut) systems are the relative extremes of the treatment environments, they are best suited to test the host preference. The means of these two treatment environments were therefore analyzed using a two-sample t-test (R, v. 3.2.2, 2015), and a significant difference between these means would signify a host preference.

To assess overall health of the systems, differences in mean gross microcosm population were analyzed using a one-way ANOVA based on treatment, and a two-way ANOVA based on treatment and block.

M. persicae showed a preference for undamaged hosts over damaged hosts. There was a significant difference ($t_{4.9}$ = 3.04, p = 0.03; Figure 1) between the undamaged-damaged and damaged-undamaged systems in the distribution of *M. persicae* between the starting and adjacent plants. In both cases, the undamaged plant hosted a higher percentage of the gross microcosm population on day 12.

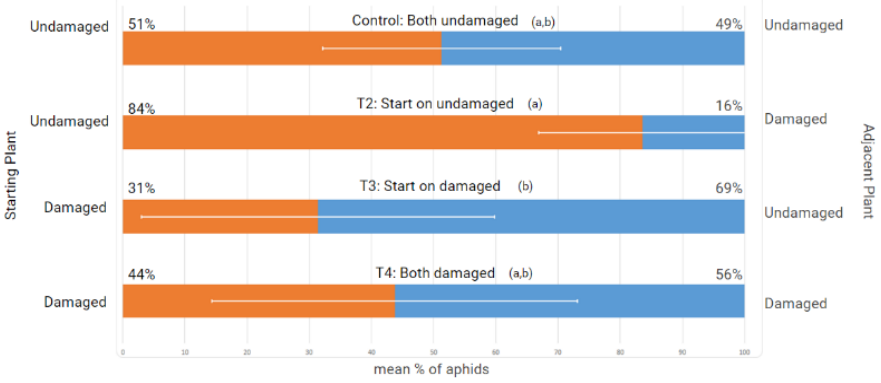


Figure 1: The distribution of a population of *M. persicae* between two *A. thaliana* plants 12 days after inoculation based on mechanical damage. Mean percentages (± 1 S.D.) of the gross microcosm population on both the inoculation plant and adjacent plant are shown for four different mechanical damage treatments. Statistically significant differences (p < 0.05) were determined by pair-wise two-sample t-tests and are indicated by different lowercase letters above bars for each system

There was no significant effect ($t_{11.5}$ = 0.40, p = 0.70) due to spatial block (Figure 2). There was also no significant relationship found in the ANOVA between the gross microcosm populations and the treatment ($F_{3, 11}$ = 0.6, p = 0.60) gross microcosm populations and block ($t_{10.1}$ = 1.28, p = 0.23); and the gross microcosm population and the interaction of treatment and block ($F_{1, 11}$ = 0.07, p = 0.80).

We noted that there were winged (alate) *M. persicae* found at multiple instances throughout the 12 day period, the first of which was found on day 4. The highest occurrences of alates were on days 6 and 7. Alates were initially found on the mechanically damaged plants, indicating that they were in a stressful environment and were attempting to leave that plant to find a better habitat.

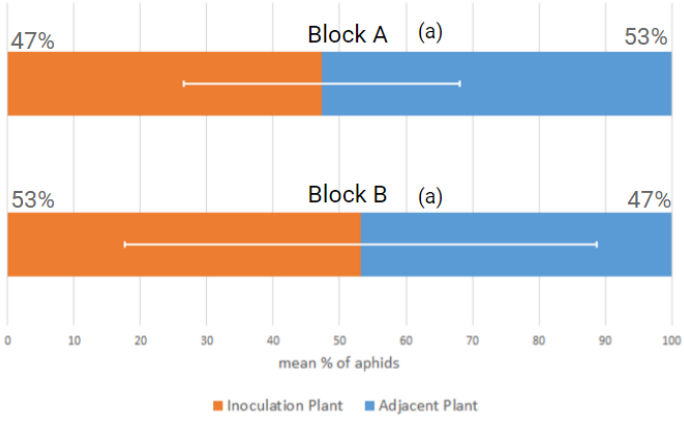


Figure 2: The distribution of a population of *M. persicae* across a two-plant *A. thaliana* system 12 days after inoculation based on spatial block. Mean percentages (± 1 S.D.) of the gross population on the inoculation plant and adjacent plant are shown for the two spatial blocks. Statistically significant differences (p < 0.05) were determined by a two-sample t-test and are indicated by different lowercase letters above bars for each spatial block.

DISCUSSION

Based on the obtained data, we determined that the *M. persicae* populations had a preference for healthy plants, the general trend showing that they moved towards the healthy plants and away from mechanically wounded plants. This demonstrates a correlation between the presence of JA and the relocation of *M. persicae* population, potentially as a result of environmentally regulated stressors. According to past research, the development of alates is an indication of a stressful host-plant environment (Sutherland and Mittler, 1971). The presence of alates in our experiment therefore implies that the treatments triggered stress in the populations. However, since our experiment was contained within two large containers each containing several plant microcosms, it is impossible to determine where an alate may have migrated to upon flight. Therefore, we cannot make any additional conclusions about our experiment based on the number of alates observed in each microcosm.

The two-plant model represents a closed small-scale ecosystem with a small number of variable parameters, and considers only one of several interactions. However, it is important to realize that host selection in aphids is a vastly complex group behaviour influenced by a variety of factors, such as aphid predation, which were not

considered in this particular study. Thus, while we identified an isolated factor for host preference, it does not demonstrate the importance of this factor relative to other factors that affect host preference, or whether this preference can be practically manifested in the realized niche of *M. persicae*.

POSSIBLE CONFOUNDING FACTORS

It was important that our results also showed that the gross populations were not significantly affected by the present factors, which would imply that the observed differences in plant populations were more heavily due to selection instead of host viability. Another important consideration must be made in relation to plant communication. Mechanical wounding triggers the release of volatiles, and one could argue that the volatile release caused by the JA pathway induced a plant defence response in neighbouring undamaged plants. However, it was shown by Thaler, et al. in 2001 that the magnitude of the plant defence response (i.e. activity of proteinase inhibitors and polyphenol oxidase) varies with the dose of JA, and so it can be reasonably assumed that the magnitude of the defence response of directly damaged plants was much higher than the indirectly volatile-induced response.

APPLICATIONS

Biocontrol over crops is a potential application of crop pest research involving *M. persicae*. Mechanical damage to crop plants is not a practical solution to preventing herbivory; however, research suggests that when JA is added to *A. thaliana* plants it increases trichome density, just as it would if the plant tissue was actually damaged (Traw and Bergelson, 2003). Thus, if the JA pathway is induced, aphid populations will try to relocate to another host due to increased defences. In an agricultural setting, JA could be applied in order to reduce aphid herbivory in crop fields, and divert their efforts towards alternate host plants surrounding the food crop. Though a large inventory of insecticides already exists, many have been shown to be problematic regarding environmental safety. JA could be less damaging to the surrounding environment, while facilitating the same defence against herbivory. Relevant research includes a study which compared JA as a deterrent spray, and conventional insecticide in relation to chilli plant growth (Awang, et al., 2015). The results showed significant evidence that JA lowers the severity and frequency of plant diseases, and improves the growth and yield of the chilli in most circumstances. Based on our results, it is probable that this effect is at least in part due to the mitigation of aphid feeding, since *M. persicae* is known to damage the chilli crop through phloem sucking and the transferral of viruses (Chakravarthy, 2015). It is important, however, to consider the variability of plant growth across plant biodiversity. Although JA was an effective alternative insect deterrent for chilli, it may not be a perfect fit for other crops. Future research efforts should explore the advantages and disadvantages of using JA as an insect deterrent on other species of food crops.

COEVOLUTION AND ADAPTATION

The difference in adaptations to the JA pathway and the salicylic acid (SA) pathway in *M. persicae* provides interesting insight into the cost-benefit of adaptations, and strategies taken to mitigate plant defence. The phloem-sucking behaviour of *M. persicae* is known to trigger the SA defence pathway; thus, for all of the plants in our experiment, this pathway would have been induced. Since the aphids showed a preference for undamaged plants, this indicates that they have adapted to tolerate the SA pathway, but not the JA pathway. Highly co-evolved tolerance mechanisms include the production of detoxification enzymes to combat allelopathic chemicals of the SA pathway (Francis, et al., 2006). However, since aphids have not adapted to tolerate the JA pathway, the concentration of JA becomes a population-regulating factor that must be taken into account when selecting a host. From a cost-benefit perspective, it is much more beneficial for *M. persicae* to produce enzymes in response to the SA pathway, as it is always induced during feeding. In contrast, it is not as beneficial for *M. persicae* to produce defences against the JA pathway since it is only induced by external factors that may not always be present, and aphids have the option of relocating to a plant with a lower concentration of JA. Future research should consider this cost-

benefit analysis when incorporating defence pathways into different experimental models. Based on the large uncertainty in our experiment, this research should also use large sample sizes to avoid population bias and collect more representative data.

CONCLUSION

In this study, we present evidence that the host selection and local spatial distribution of *M. persicae* is influenced by the mechanically damaged state of *A. thaliana*, showing a preference for undamaged host plants. This behaviour is likely an adaptation to the significant increase in plant defences from the induced JA pathway. The magnitude of this preference relative to other factors affecting host preference is not yet known, and the interaction of factors affecting host preference continues to be an area of interest.

MATERIALS AND METHODS

STUDY SYSTEM

The mobile organism we selected for this experiment was the *Arabidopsis*-reared green peach aphid, *M. persicae*. Although thought to originate from Asia, *M. persicae* currently has populations widespread across the globe. *M. persicae* is considered an economically important crop pest and an important distributor of various plant viruses (Bass, et al., 2014; Li, et al., 2015). It has a wide host range, including important crops such as potato and members of the Brassicacea family, totalling over 400 specific plant species from approximately 50 families (Bass, et al., 2014; Louis and Shah, 2013; Li, et al., 2015). *M. persicae* has a short development period and a relatively short lifespan, lasting approximately 50 days from the early instar phase to death (Musa, et al., 2004). For these reasons, studies regarding aphid host selection and aphid population control are of high practical relevance.

The host plant we selected for this study was wild-type *A. thaliana* (accession line: Col). Although not an economically relevant plant, this herbaceous member of the Brassicacea family is often used as a model organism in scientific studies. As such, much is known about *A. thaliana*'s fast life cycle, easily manipulated genome, and herbivore defence pathways (Mewis, et al., 2006). Its short generation time, small size, and inclusion in the *M. persicae* host range make it an ideal host plant for this study.

TREATMENTS

In order to determine the effects of host-plant health and condition on aphid host selection, we performed a microcosm experiment using potted wild-type *A. thaliana* and *Arabidopsis*-reared *M. persicae* from a laboratory colony. There were four treatments of microcosms, each consisting of a pair of *A. thaliana* plants connected by two "bridges" made from thin string and wooden dowels (Figure 3). Each microcosm consisted of three adult or late instar *M. persicae* inoculated near the bottom of the central stem on Plant 1 in each system. Each treatment had four replicates for a total of 16 microcosms used in the experiment. The microcosms were stored on large plastic trays which were covered with clear plastic domes with mesh tops for ventilation. Each tray held two replicates of each treatment. The treatments were arranged randomly on the trays.

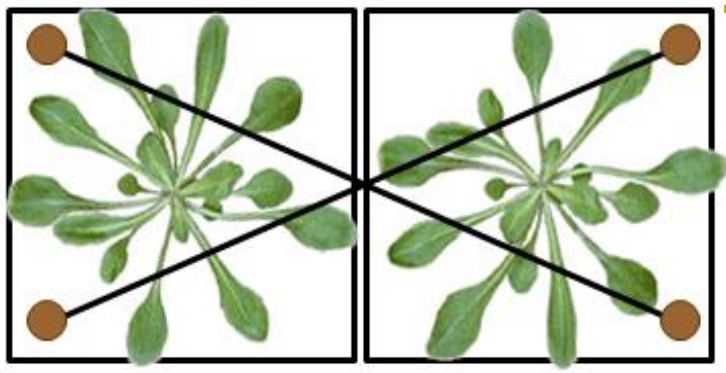


Figure 3: The setup of a two-plant microcosm studied in this experiment. Each microcosm consisted of two adjacent *A. thaliana* plants varying in state of mechanical damage, and three *M. persicae* were inoculated on one of the plants. String bridges were erected with wooden dowels to facilitate aphid movement.

The four treatments, which will be described as starting plant-adjacent plant, included: (1) an undamaged-undamaged system, which acted as the control, (2) an undamaged-damaged system; (3) a damaged-undamaged system; and (4) a damaged-damaged system, which acted as the negative control. For each respective plant that was damaged, 30% of the leaves were cut to ensure

plant survival. Three cuts were made to each cut leaf using scissors: one parallel to the main vein of the leaf at the leaf tip, and two perpendicular to the main vein with one on either side of the leaf. No cuts were made to remove sections of the leaf. We cut the leaves such that approximately an equal number of leaves were damaged in each quadrant of the plant.

The experiment lasted 12 days. We recorded aphid count on each plant in each microcosm on 9 of the 12 days the experiment was run (days 1, 4-8, 11, 12). All microcosms were watered once on day 7. In order to remove bias from the experiment, all microcosms were exposed to the same duration of sunlight and darkness, and the same environmental conditions.

ACKNOWLEDGEMENTS

This research was adapted from work done for the Plant-Animal Interactions module of the Integrated Science program. We thank McMaster University and the Integrated Science program for access to laboratory materials, as well as Russ Ellis, Dr. Chad Harvey and Sebastian Irazuzta for their helpful feedback and support. We would also like to thank you, the reader, for your time. Here is a haiku for you:

The Big Aphid: A Haiku

Big aphid leaves home
Bridge of string to distant land
Life without JA

AUTHOR CONTRIBUTIONS

All authors contributed equally to the final paper in terms of editing, writing, data collection, and experimentation. However, it should be noted that M.C. had a primary focus on statistical analysis, J.P. had a primary focus on experimental set-up, and C.A., R.A., and R.E. had a primary focus on preliminary research and analysis of observations.

REFERENCES

Awang, N.A.A., Ismail, M.R., Omar, D., and Islam, M.R., 2015. Comparative study of the application of jasmonic acid and pesticide in chilli: effects on physiological activities yield and viruses control. *Institute of Tropical Agriculture,* 31(3), pp.671-681.

Bass, C., Puinean, A.M., Zimmer, C.T., Denholm, I., Field, L.M., Foster, S.P., Gutbrod, O., Nauen, R., Slater, R., and Williamson, M.S., 2014. The evolution of insecticide resistance in the peach potato aphid, *Myzus persicae. Insect Biochemistry and Molecular Biology,* 51, pp.41-51.

Blackman, R.L., 1974. Life-cycle variation of *Myzus persicae* (Sulz.) (Hom., Aphididae) in different parts of the world, in relation to genotype and environment. *Bulletin of Entomological Research,* 63, pp.595-607.

Chakravarthy, A.K., 2015. *New horizons in insect science: towards sustainable pest management.* Karnataka, India: Springer.

De Vos, M., Kim, J. and Jander, G. 2007. Biochemistry and molecular biology of *Arabidopsis*–aphid interactions. *Bioessays,* 29(9), pp.871-883.

Francis, F., Gerkens, P., Harmel, N., Mazzucchelli, G., de Pauw, E. and E. Haubruge, 2006. Proteomics in *Myzus persicae*: effect of aphid host plant switch. *Insect Biochemistry and Molecular Biology,* 36, pp.219–227.

Goggin, F.L., 2007. Plant–aphid interactions: molecular and ecological perspectives. *Current Opinion in Plant Biology,* 10, pp.399–408.

He, Y., Fukushige, H., Hildebrand, D. and Gan, S. ,2002. Evidence supporting a role of jasmonic acid in *Arabidopsis* leaf senescence. *Plant Physiology,* 128(3), pp.876-884.

Jaouannet, M., Rodriguez, P., Thorpe, P., Lenoir, C., MacLeod, R., Escudero-Martinez, C. and Bos, J. 2014. Plant immunity in plant-aphid interactions. *Frontiers in Plant Science,* 5.

Kettles, G. J., Drurey, C., Schoonbeek, H., Maule, A.J. and Hogenhout, S.A., 2013. Resistance of *Arabidopsis thaliana* to the green peach aphid, *Myzus persicae,* involves camalexin and is regulated by microRNAs. *The New Phytologist,* 198, pp.1178-1190.

Li, J., Cao, J., Niu, J., Liu, X. and Zhang, Q., 2015. Identification of the population structure of *Myzus persicae* (Hemiptera: Aphididae) on peach trees in China using microsatellites. *Journal of Insect Science,* 15(1), p.73.

Louis, J. and Shah, J., 2013. *Arabidopsis thaliana—Myzus persicae* interaction: shaping the understanding of plant defence against phloem-feeding aphids. *Frontiers in Plant Science,* 4, p.213.

Mewis, I., Tokuhisa, J.G., Schultz, J.C., Appel, H.M., Ulrichs, C. and Gershenzon, J., 2006. Gene expression and glucosinolate accumulation in *Arabidopsis thaliana* in response to generalist and specialist herbivores of different feeding guilds and the role of defence signaling pathways. *Phytochemistry,* 67, pp.2450-2462.

Musa, F.M., Carli, C., Susuri, L.R. and Pireava, I.M. ,2004. Monitoring of *Myzus persicae* (Sulzer) in potato fields in Kosovo. *Acta Agriculturae Slovenica,* 83, pp.379–385.

Patrick J. and Moran, G., 2001. Molecular responses to aphid feeding in *Arabidopsis* in relation to plant defence pathways. *Plant Physiology,* 125(2), p.1074.

R Development Core Team, 2014. *R: a language and environment for statistical computing.* R Foundation for Statistical Computing.

Rockwood, L. 2015. *Introduction to population ecology,* 2nd ed. West Sussex, United Kingdom: John Wiley & Sons.

Seigler, D. S. 2012. *Plant secondary metabolism.* New York : Springer Science & Business Media.

Sutherland, O.R.W. and Mittler, T.E., 1971. Influence of diet composition and crowding on wing production by the aphid *Myzus persicae. Journal of Insect Physiology,* 17, pp.321-328.

Thaler, J., Stout, M., Karban, R. and Duffey, S. 2001. Jasmonate-mediated induced plant resistance affects a community of herbivores. *Ecological Entomology,* 26(3), pp. 312-324.

Traw, M.B. and Bergelson, J., 2003. Interactive effects of jasmonic acid, salicylic acid, and gibberellin on induction of trichomes in *Arabidopsis. Plant Physiology,* 133, pp.1367–1375.

Vargas, R.R., Troncoso, A.J., Tapia, D.H., Olivares-Donoso, R. and Niemeyer, H., 2005. Behavioural differences during host selection between alate *virginoparae* of generalist and tobacco-specialist *Myzus persicae. The Netherlands Entomological Society,* 116, pp.43-52.

War, A.R., Hussain, B. and Sharma, H.C., 2013. Induced resistance in groundnut by jasmonic acid and salicylic acid through alteration of trichome density and oviposition by *Helicoverpa armigera* (Lepidoptera: Noctuidae). *AoB Plants,* 5, p.53.

Comparison of Mathematical Models of Opinion Dynamics

AURORA BASINSKI-FERRIS
Integrated Science Program, Class of 2018, McMaster University

SUMMARY

Sociophysics is an emerging interdisciplinary research field that uses theories and methods developed by physicists as a modelling framework for various social science phenomena. This paper presents the Sznajd model, which is the basis of one of the simplest yet most efficient methods to predict the collective human behaviour in a binary opinion society. This model is a variation of the Ising model – a mathematical model used in statistical mechanics when studying the ferromagnetic phase transitions for the magnetization of a system based on the interaction between the neighbouring atomic spins of the system. In both the Ising model and the Sznajd model, individuals are only allowed to have a binary opinion – yes (spin up) or no (spin down). The Sznajd model is based on the idea of social validation; the model requires that two neighbours share the same opinion to be able to convince individuals around them of their opinion. Using MATLAB, this paper applies a two-dimensional simplified version of the Sznajd model to track the opinion movement over many time steps in a lattice of individuals with an initial randomly generated distribution of opinions. The results of this simplified model are then analyzed using psychological principles to examine the validity of the outcomes.

Received: 11/13/2016 **Accepted:** 03/07/2017 **Published:** 07/31/2017
URL: https://journals.mcmaster.ca/iScientist/article/view/1357

Keywords: sociophysics, Sznajd Model, opinion dynamics, mathematical modelling, MATLAB

INTRODUCTION

The idea of applying concepts from the natural sciences to try to explain social science phenomena is at least 25 centuries old. One of the first recorded comparisons between the properties of natural sciences and the behaviour of humans occurred when the Greek philosopher Empedocles stated an idea to the effect of "Humans are like liquids: some mix easily like wine and water, and others like oil and water refuse to mix." (Stauffer, 2009). Since then, there have been a few individuals who drew comparisons between ordered physical systems and the behaviour of populations. For example, in 1942, Majorana compared quantum physics to social probabilities in human behaviour. However, with the exception of a select few individuals, the field of sociophysics took off in the early 2000s with the popularization of ideas such as opinion dynamics within academia (Stauffer, 2013).

Mathematical models of opinion formation examine society as a physical system in which individuals and their interactions are viewed as the microscopic scale of the system governed by rigid rule sets, and the emergent macroscopic trends are used to reflect overall societal opinion. In general, these models deal with binary opinion systems in which there is a general assumption that the opinion of an individual is in some way related to the opinions of their neighbours (Oz, 2008).

Models of opinion dynamics, and sociophysics in general, rely on the law of large numbers. This law is a statistical theorem that states that as the number of randomly generated variables increases, the average approaches the theoretical mean (Routledge, 2015). A simple system used to understand this is coin flipping. When just one coin is flipped, although the probability of either outcome is 0.5, you cannot predict with any certainty whether that single coin toss will show a head or tails. However, after 500 coin tosses, it is

reasonable to assume that the long-term behaviour is approaching the 0.5 probability. Thus, this law applied to opinion dynamics means that with a large population, individual fluctuations are averaged out, and general trends in data are more visible. In turn, this also implies that the justification for the human aspect of these models stems from mass psychology rather than individual psychology.

This paper offers a brief history of the models developed to deal with the opinions of populations. However, the focus will lean more towards the examination of two variations of a fairly recent model of opinion dynamics called the Sznajd model. These two model variations were simplified and represented in MATLAB to analyze and compare the behaviour and validity of these models given an initial random population with varying size.

MATHEMATICAL BACKGROUND

Many models have been proposed to try to examine the opinion dynamics of populations. In general, these follow the theme of an initially disordered system that is brought to a more ordered state under certain conditions (Castellano, Fortunato, and Loreto, 2009). The key part of opinion dynamic models is the overarching rule that guides how interactions between various individuals in the population occur. In most models, the opinion of a single individual is updated after each time step; this change is affected by the opinions of the individual's neighbours. Additionally, all models must define the space in which their population exists. The models presented in this paper treat populations as discrete points on a lattice rather than a continuous plane (Sen and Chakrabarti, 2013). This setup results in an assumption that individuals in the model are stationary.

ISING MODEL AND VARIATIONS

In physics, the 1924 Ising model is a mathematical model used for ferromagnetism (Sznajd-Weron, 2005). The Ising model looks at each site on a square lattice as an electron with a value assigned of either +1 (spin up) or -1 (spin down). Thus, each pair of neighbours at <i,j> has an energy of -JSiSj where J is some proportionality constant and S denotes spin. This means that each parallel pair of spins contributes -J to the total energy, and each antiparallel pair contributes +J to the energy (Stauffer, 2013).

Therefore, the total energy of the system is given by:

$$H = -J \sum_{<i,j>} s_i s_j$$

To minimize energy, each spin is pushed to be aligned with its neighbours (Castellano, Fortunato and Loreto, 2009). Thus, the Ising model typically results in a higher probability of two neighbouring electrons having the same spin (Stauffer, 2009). Metropolis dynamics of the Ising model takes each elementary move to be a single spin change. This change in spin is accepted with a probability of:

$$P = e^{-\Delta E / k_B T}$$

where T is temperature, k_B is the Boltzmann constant, and ΔE is the change in energy when this change takes place (Castellano, Fortunato, and Loreto, 2009). As this probability of acceptance is dependent on temperature, there is a critical temperature that denotes whether or not macroscopic ordering will occur. Above the critical temperature, there is macroscopic disorder even over many time steps. However, despite the disorder, there are still local clusters of just spin up (+1) or just spin down (-1). Below the critical temperature, macroscopic ordering occurs, because these local clusters spread to become global (although limited by the bounds of the lattice). Thus, the stable states of either all positive or all negative spins would emerge (Castellano, Fortunato, and Loreto, 2009).

The dynamics that emerge from the Ising model due to physical interactions can be applied to a human system to develop a model of opinion dynamics (Stauffer, 2009). In this case, neighbours on a lattice influence each other such that adjacent individuals 'want' to have similar opinions because energy functions as 'unhappiness' due to disagreement. Due to the application of physical phenomena to social dynamics, temperature has a social meaning. The social application of temperature denotes both the overall approximation of events that influence decision making of individuals not included in the microscopic rules and denotes the tolerance of the individuals in the population. As expected, changing these factors can change the behaviour of

the system, just as going above or below critical temperature will change the ferromagnetic system. When temperature can grow arbitrarily large, neighbours will not affect an individual's views. When temperature is of an intermediate value, small clusters form with similar opinions, but no overarching domains are present. When temperature is small, these domains extend to infinite size and all members of the population have the same opinion (Stauffer, 2009).

SZNAJD MODEL

Since the early 2000s, the vast majority of investigations into opinion dynamic models have been centered on one of three models: Sznajd, Krause-Hegselmann, and Deffuant (Stauffer, 2009). Sznajd is addressed in this paper as it fits the best with old investigations, such as Ising variants in that it utilizes discrete opinions. In contrast, the Krause-Hegselmann and Deffuant models allow agents in their population to have continuous opinions within a range (Stauffer, 2009).

The Sznajd model uses the concept of Ising spins, but slightly modifies the physical model to fit with human interactions rather than to make a direct correlation with physical phenomena. It is based around modifications of the microscopic interaction rules to make them more in line with human group psychology (Sznajd-Weron, 2005). The main relation between the Sznajd and Ising models is the application of the Ising spin chain idea, in which the spins are defined as binary variables that can influence the 'opinion' of neighbours (Sznajd-Weron and Sznajd, 2000). The basic rule of Sznajd states that, if there is a pair of neighbours in agreement, the individuals surrounding them will change their opinions to be in line with this spin. Thus, in one dimension, it can be stated that if Si=Si+1 then Si-1=Si and Si+2=Si. Correspondingly, in two dimensions, a pair of neighbours in agreement will influence the six neighbours surrounding them (Sznajd-Weron, 2005).

There are two variations of the Sznajd model. Both variations deal with the agreement of neighbours in the same way, but they differ in their treatment of the disagreement of neighbours (Sznajd-Weron and Sznajd, 2000). In the original model put forward in 2000, if Person A and Person B are neighbours in disagreement, the surrounding individuals of Person A take the opinion of Person B and vice versa. This was meant to emulate the act of an argument whereby neighbours are influenced. The long-term behaviour of a system following this dynamic rule would either be ferromagnetic (consensus) or antiferromagnetic (stalemate). For this reason, the Sznajd model was originally named "United we stand, Divided we fall" (Sznajd-Weron and Sznajd, 2000). The second variation of the Sznajd model was proposed in 2005. In this model, if Person A and Person B are neighbours in disagreement, the surrounding individuals of Person A would take Person A's opinion and vice versa. The only stable state that can come of this model is ferromagnetic (consensus) (Sznajd-Weron, 2005).

There are a few key psychological phenomena that contribute to the validity of the Sznajd model as a simplified description of human behaviour. The primary focus of the dynamic rules is on social validation. Simply, social validation is the phenomenon in which individuals tend to conform to the actions of others when they are in a group. In terms of group psychology, this means that when more than one person has a certain behaviour, individuals around them will start to adopt this behaviour (DeMers, 2015). This phenomenon was demonstrated through Asch's conformity experiments in the 1950s which tested the extent to which people tend to conform to the majority (Bond and Smith, 1996). The experimental setup involved showing a group of lines to test subjects. Before the subjects answered if the lines were of the same length, experimenters acting as subjects would incorrectly state an answer. Results found that subjects of the experiment would usually conform to this false opinion; only 29% of the subjects stated the correct answer rather than the most popular answer (Bryn Mawr College, 2000). Asch concluded that the perception of group consensus results in the conformity of individuals to a set opinion. Finally, it was found that an individual was more likely to conform to popular opinion if the individual identified with the majority – that is, if there are similar characteristics that both the individual and the majority hold (Bond and Smith, 1996). Thus, in the Sznajd model, this rule applies due to the 'majority' opinion of the pair affecting its neighbours rather than distant individuals.

METHODS

The modified MATLAB models created for this paper were designed to compare the two variations of the Sznadj model. Each variation consists of two sections of code – the first examines the behaviour of the system over a set number of time steps, while the second examines the number of time steps needed to reach a stable state (The MathWorks, Inc, 2015).

All initial matrices used for the two models for each variant were developed using the random matrix function. This automatic function randomizes a matrix of a given size with numbers in the range 0 to 1. This was converted to a random matrix of discrete values of -1 and +1, by setting original values between 0 and 0.5 to equal -1, and values between 0.5 and 1 to equal 1.

Both models apply their respective dynamic rules by utilizing a concept known as a Monte Carlo simulation. Monte Carlo is the act of approximating an overall expectation based on the behaviour of random samples; as samples are chosen by chance, this technique was named after an international gambling city (Anderson, 1999). When performing a mathematical model on a system where all the points in that system should behave according to a predefined rule, such as the Sznajd Model, a Monte Carlo simulation can be used. The Monte Carlo method takes random samples, and applies the rule at that site (Anderson, 1999).

In these models, the Monte Carlo simulations are performed using the randomization functions of which MATLAB is capable. First, an initial point (i,j) is chosen at random by generating two separate random numbers that fall in the matrix dimensions for both i and j. As the system is a two-dimensional matrix, each point has four neighbours. Thus, after this initial point is generated, another number is randomly generated between 1 and 4. This number chooses the pair that the dynamic rule will be applied to – a pair composed of that initial point and a randomly chosen neighbour.

As discussed in background, the two variations addressed have the same rule for when the pair chosen are in agreement, but deal with disagreement in different manners. Thus, in both variants, when the randomly chosen pair have consistent opinions, the six neighbours surrounding the pair take the opinion of this pair (Figure 1).

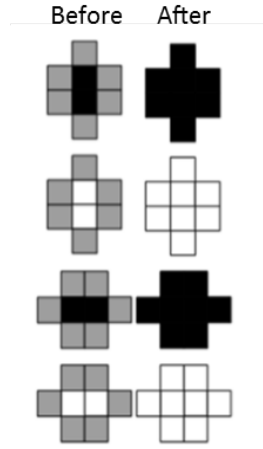


Figure 1: Illustration of the cases of initial pairs in agreement on the lattice generated by the MATLAB model. The grey boxes in the "Before" column indicate that those boxes could take either white or black values and it would not affect the outcome. At each time step, the initial pair would be randomly chosen, and the "After" column displays the behaviour of the neighbours near the pair.

The two variants of the Sznajd model that were compared differ in their treatment of disagreement. The microscopic description of this behaviour is shown in Figure 2.

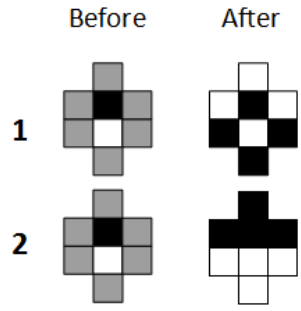


Figure 2: Illustration of the cases of initial pairs in disagreement on the lattice generated by the MATLAB model. The grey boxes in the "Before" column indicate that those boxes could take either white or black values and it would not affect the outcome. At each time step, the initial pair would be randomly chosen, and the "After" column displays the behaviour of the neighbours near the pair as per the rule of the model variant (denoted 1 and 2 in the row headings).

When designing the modelled system, periodic boundary conditions were utilized. This was done to avoid errors that occur at boundaries if the general interior rule was applied universally. For example, if looking at a pair of neighbours composed of a random point and a neighbour to the right, one of the neighbours that should be affected is one to the left of the initial point. If away from the boundaries, this could be stated as:

If A(i,j) is equal to A(i,j+1)

Then A(i,j-1) should equal A(i,j)

However, with a lattice of finite dimensions, the problem is encountered where the initial point could be on a boundary. For example, if the initial point is in initial column 1, a point one column to the left (A(i,j-1)) does not exist. Thus, the modular function is implemented to make assigning a point which should lie outside the matrix dimensions to loop back around to the other side of the matrix.

For example, take A(i,1) as the randomly generated point in a lattice that is 10 by 10. If we use the general 'then' statement as A(i,mod(j-2),y)+1), then we would use A(i, mod(-1,10) +1). This outputs the resultant point of A(i,10) meaning that the point 'to the left of the 1st column' is equal to the last column. Importantly, this modular design still outputs what we would expect when the initial point is not on the boundaries. If we take the point A(i,6) as the randomly generated point in a 10 by 10 lattice, we can see that applying this modular 'then' statement would result in A(i,mod(4,10)+1), which is equal to A(i,5) as expected. When applied in all directions, the use of the modular function results in the system looping and essentially functioning as a sphere.

For each model variation, there were two different sections of code developed, which each output different information. The two sections of code for each variation follow the descriptions included below; they were similar in that they all allowed for varied matrix size, followed the same dynamic rules for respective variation, and utilized the modular function to develop the space on which the individuals reside as a sphere.

RESULTS

The first section of code for each rule was relatively simple. It was made such that it applied the dynamic rules for a given number of time steps, showed the movement of the matrix over time, and graphed the percent of positive (yes) opinions on the lattice. Tracking the movement of the matrix over time utilized a three dimensional array indexed by time step. At each time step in this array, one could find the two-dimensional matrix that displayed the state of the population at that time.

The second section of code for each model variant extracted data useful for effectively comparing the two variants. It calculated the number of steps required by each variant to reach one of the stable states discussed in the background – stalemate or consensus. Interestingly, each of those options has two types; intuitively, it is known that consensus can occur either as all yes (+1 spin) or all no (-1 spin). Stalemate can also occur in two ways, because the tiles alternating can be shifted – such that the top left corner is yes, or such that the top left corner is no. For each variant, this model computed the number of steps required for a set number of trials on a set number of matrices. The output was a scatter plot of the average of the 20 trials for each matrix, an overall average across all trials and matrices, and a pie chart showing the proportion of the stable states achieved that were of each type (Figure 3).

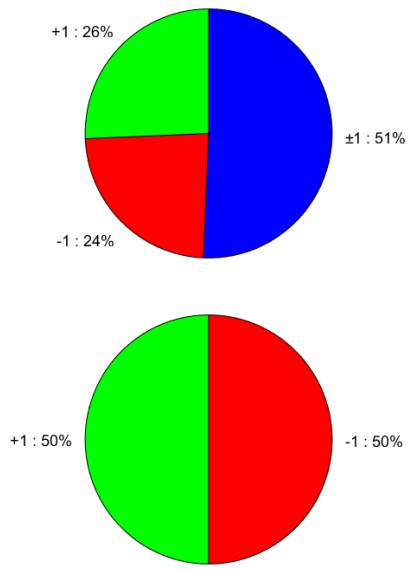


Figure 3: The proportion of each type of stable state achieved after 30 trials of 30 randomly generated 12 by 12 matrices. (A) displays the output for variation 1, described in the methods, while (B) shows the outputs for variation 2.

Examination of the general results from the second model of each rule shows one of the clearest distinctions between the two dynamic rules. Rule 1, which functions using the idea of argument when two neighbours disagree, can yield both a consensus and stalemate stable state. On the other hand, rule 2 yields consensus unanimously. However, in both of these models, there doesn't appear to be a trend showing that any one possible stable state is preferred over the others.

When 30 different random matrices are tested with 30 trials each (net of 900 trials), there appears to be around equal partition of the stable states possible. For rule 1, there are 4 stable states of which two were stalemate, and two were consensus (yes or no). As seen in Figure 3, half of trials yielded a stalemate stable state, while a quarter yielded each of yes consensus and no consensus. For rule 2, in which there are two stable states, half of the trials yielded a yes consensus and half of the trials yielded a no consensus.

Thus, though the different variants have different possible stable states, all stable states are essentially equally possible for each variant. This means that there is no inherent pressure towards a final outcome within each model. However, the second variant has no possibility of stalemate and the population present always reaches consensus.

As shown in Figure 4, both dynamic rules result in a wide range between the average of 30 trials for each matrix. However, it becomes clear (especially when looking at very high numbers of trials) that both the average number of runs across all trials of all matrices, and the range for the average of 30 trials, vary when the different dynamic rules are applied. The first dynamic rule, which utilizes the idea of an 'argument' when the chosen pair disagrees, results in a mean of 1.4033*103 iterations necessary to reach a stable state. Alternatively, the second dynamic rule resulted in a mean of 2.4618*103 iterations until a stable state. This trend between the two variants remained mostly consistent over different matrix sizes and trial numbers.

There are a few possible reasons that could account for this difference in number of iterations required to reach stable state. The first possible reason is fairly simple – the first variation is able to reach either stalemate or consensus. Thus, the first variation has four possible states it can reach in order to achieve stability. Alternatively, the second variation can only reach consensus; thus, it only has two states that it can reach in order to be stable.

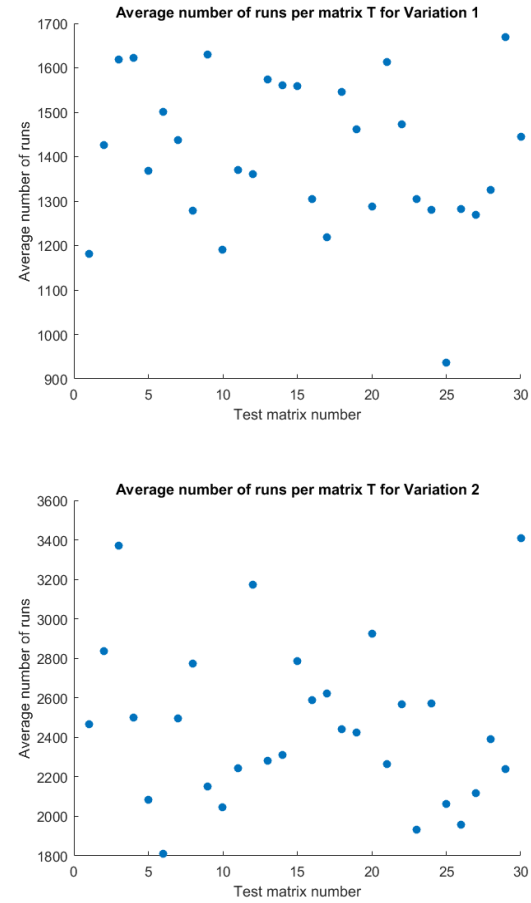


Figure 4: Scatter plots of the average of number of runs to reach a stable state. This data was generated using 30 trials for 30 separate randomly generated initial 12 by 12 matrices. (A) shows the data generated when dynamic rule 1 was applied while (B) shows the data generated when dynamic rule 2 was applied.

Another possibility is the interaction between the various rules and states. In the case of the second rule, the black and white solid sections tend to get more separated and have more stability on their side of the lattice. On the other hand, there is more interaction and less stability with the first rule. Often dynamics such as that displayed in Figure 5 occur.

Figure 5 shows the dynamic interaction between the checkered stalemate state and the consensus (white) state. There were very few areas with no movement in either section due to the dynamic rules, and one state can quite quickly take over the matrix. This can be seen more quantitatively in Figure 6. This figure shows there is a large

fluctuation in the percent of "yes" voters on the matrix irrespective of the stable state eventually reached. The spikes up and down, which can encompass a large range, show that there is quite a bit of movement between various areas displaying a stable state behaviour.

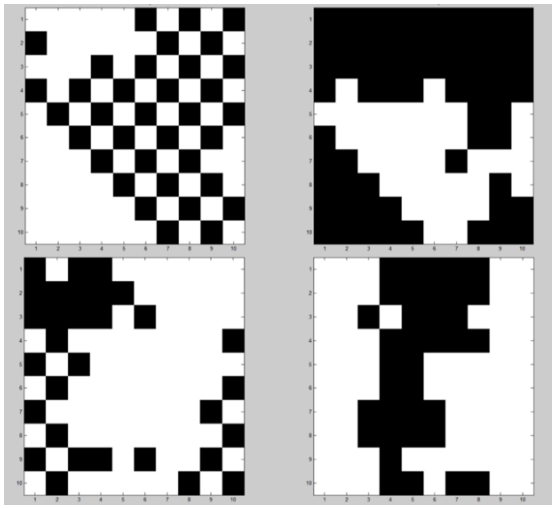


Figure 5: Still frames throughout the evolution of two randomly generated 12-by-12 matrices. Each row shows the two iterative rule variations applied to the same initial matrix. The left column shows two instances of variation 1, while the right column shows variation 2.

DISCUSSION

Models inspired by physical phenomena can work effectively when applied to sociophysical areas such as opinion dynamics. The simplification presented in this paper was developed using the dynamic rules proposed in the established Sznajd model which uses concepts from the Ising model of ferromagnetism and modifies them to better fit with psychological phenomena. This paper presented two different variations of the Sznajd model that each used distinct iterative rules. The setup for both models required a Monte Carlo simulation of a lattice. The two variations used the same rule when a pair of agreeing neighbours was chosen, whereby the six surrounding neighbours took that opinion. In the case of disagreement, the first variation had the neighbours of individual A take the opinion of individual B; the second variation had the neighbours of individual B take the opinion of individual A.

The results yielded were quite different for the two variations over many time steps. When these iterative rules were applied on the same initial setup, the first variation resulted in a long-term stable state of either consensus or stalemate. Alternatively, the second variation resulted exclusively in consensus.

The accuracy of both of these stable states to psychological phenomena can be argued. Due to the Asch conformity experiment on which this model is based, the idea of long-term consensus has merit. If individuals generally wish to fit in with the opinions of others and adhere to the majority, consensus seems possible in many situations. Although possibly anecdotal, examples of population conformity throughout history further add to our perception of the validity of achieving consensus in a population. However, especially if regarding points on the lattice as small groups of likeminded people rather than individuals, the stalemate solution cannot be discounted completely. Although we do not traditionally think of individuals as arranging themselves in such a way, the idea of small

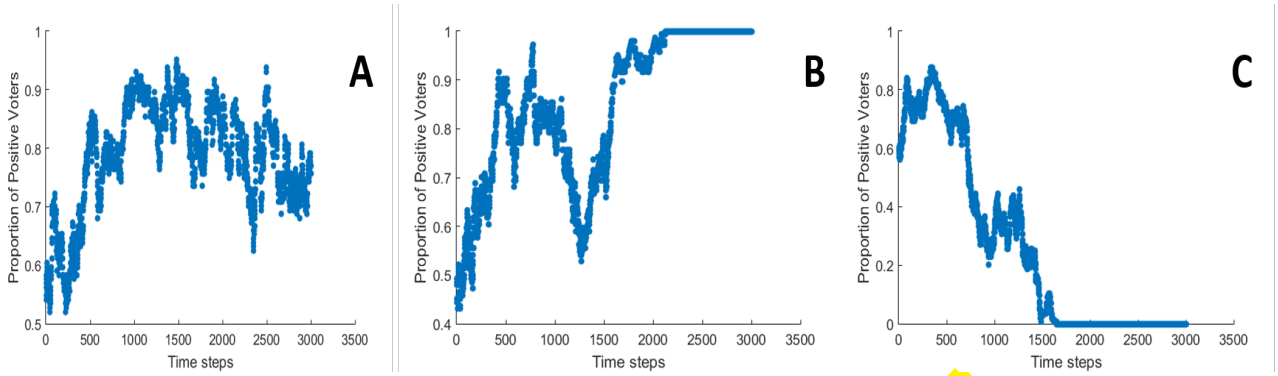


Figure 6: Graphs showing the percent of positive voters over iterations when running variant 1. (A) corresponds to an eventual stalemate. (B) corresponds to positive consensus. (C) corresponds to negative consensus. All graphs show large fluctuations in the proportion of positive voters throughout their respective evolutions to stable states.

communities of like-minded people staying clustered and separated from those who disagree with them could apply in some situations.

CONCLUSION

The choice of model variant may depend more on how it will be applied, rather than an inherently better solution. If looking at individuals, perhaps the evolution that always achieves conformity is more fitting with group psychology. However, if trying to model a larger scale system such as clusters of people in agreement, without regarding intra-cluster dynamics, the first model presented may be of more use. It is important to note that no model can ever exactly predict the behaviour of a system of biological individuals. However, as more variables are considered and modelled, the model can approach accuracy much more closely.

ACKNOWLEDGEMENTS

The author would like to thank Dr. George Dragomir (McMaster University) for supervising this project, and providing continuous assistance throughout

REFERENCES

Anderson, E., 1999. Monte Carlo Methods and Importance Sampling. [online] University of Washington: Lecture Notes for Stat 578C. Available at: <http://ib.berkeley.edu/labs/slatkin/eriq/classes/guest_lect/mc_lecture_notes.pdf> [Accessed 24 Mar. 2016].

Bond, R. and Smith, P., 1996. Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. Psychological Bulletin, 119(1), 111-137.

Bryn Mawr College, 2000. Solomon E. Asch 1907-1996. [online] The Solomon Asch Center for Study of Ethnopolitical Conflict. Available at: <http://www.brynmawr.edu/aschcenter/about/solomon.htm> [Accessed 21 Mar. 2016].

Castellano, C., Fortunato, S. and Loreto, V., 2009. Statistical physics of social dynamics. Reviews of Modern Physics, 81(2), 591-646.

DeMers, J., 2015. The Importance Of Social Validation In Online Marketing. Forbes. [online] Available at: <http://www.forbes.com/sites/jaysondemers/2015/02/19/the-importance-of-social-validation-in-online-marketing/#4fd3ee6535c8> [Accessed 21 Mar. 2016].

Oz, M., 2008. Sznajd Model and Its Application to Politics. [online] p.11. Available at: <https://pdfs.semanticscholar.org/73ab/024a432d47ce20eaf407b995dc7356511588.pdf>.

Routledge, R., 2015. Law of large numbers. In: Encyclopædia Britannica. [online] Available at: <http://www.britannica.com/science/law-of-large-numbers> [Accessed 21 Mar. 2016].

Sen, P. and Chakrabarti, B.K., 2013. Sociophysics: An Introduction. New York: Oxford University Press.

Stauffer, D., 2009. Opinion dynamics and sociophysics. Encyclopedia of Complexity and Systems Science, 6380-6388.

Stauffer, D., 2013. A biased review of sociophysics. Journal of Statistical Physics, 151(1), 9-20.

Sznajd-Weron, K., 2005. Sznajd model and its applications. Acta Physica Polonica B, 36(8), 2537-2549.

Sznajd-Weron, K. and Sznajd, J., 2000. Opinion evolution in closed community. International Journal of Modern Physics C, 11(6), 1157-1165.

The MathWorks, Inc, 2015. MATLAB (R2015a). [computer program] The MathWorks, Inc.

Secrets in Math: How do we Fly?

SIMON ZHANG
Integrated Science Program, Class of 2018, McMaster University

SUMMARY

This article covers the history of flight from 1400s to the modern era, focusing on d'Alembert's paradox and explanations of how lift is created. It is unnerving to think that it wasn't until 2008 that scientists were finally able to answer the question of how planes remain in the air. D'Alembert's paradox has troubled mathematicians and aeronautic experts for years. It was a simple statement made in 1749, but the implications have influenced the field of fluid dynamics for centuries to come. The statement made by d'Alembert essentially concluded that there is no mathematical explanation for the physical observation of drag on any solid object moving through a fluid, making airborne flight an unexplainable enigma. However, in 2008, a paper was published finally providing a coherent resolution to d'Alembert's paradox as well as fully explaining how lift and drag are created (Hoffman and Johnson, 2008; 2009). After disproving the previous theories provided by Prandtl and Kutta-Zhukovsy, Johan Hoffman and Claes Johnson built on the work of d'Alembert and Stokes and were able to identify the instability mechanism which their predecessors overlooked. Using this mechanism, they were able to explain the cause of net drag on the wing. With this understanding, airplane developers can take greater steps to improving the design of airplane wings and make air travel much more efficient.

Received: 11/13/2016 **Accepted:** 03/10/2017 **Published:** 07/31/2017
URL: https://journals.mcmaster.ca/iScientist/article/view/1356

Keywords: flight, tensors, d'Alembert's paradox

INTRODUCTION

Humans have always been interested in conquering new frontiers. This was evident thousands of years before the modern era, in Ancient Greece; from the myths of gods overhead to studying the treks of the stars through the heavens, they were seduced by the concept of flight. Although humans have been interested in flying for thousands of years, the field of aeronautics only began in 1485 with Da Vinci's design of "The Ornithopter". However, it took until 1903 for the first flying machine to be invented by the Wright Brothers, meaning it took over 400 years to go from design to creation (NASA, 2016). That being said, the mathematics required to achieve these developments has lagged even further behind the physical advancements in the real world. This rift between the mathematics and engineering behind flight started in 1749 with the formation of d'Alembert's paradox (Claes Johnson, 2012a). Jean le Rond d'Alembert was a French mathematician who used Euler's equations to show that the drag on a body moving through in an negligibly viscous (inviscid) and incompressible fluid was zero. This did not align with any observations made in the physical world and caused the engineering branch of fluid mechanics, known now as hydraulics, to separate into its own field, different from the field of theoretical fluid mechanics. It was only recently that the mathematics finally caught up with the physical observations. This occurred in 2008 when an explanation and resolution for d'Alembert's paradox was sufficiently tested and explained by Johan Hoffman and Claes Johnson (Hoffman and Johnson, 2008; 2009). This means that for over 250 years, we had no concrete answer to the fundamental question of: "What keeps planes up?" This paper informs readers about the history of the development of the mathematics behind fluid dynamics as well as to depict how the current resolution of d'Alembert's Paradox explains the possibility of flight.

GENERAL THEORY OF FLIGHT

Though humans had been interested in flight for thousands of years beforehand, the individual who made the first documented strides towards understanding was Leonardo da Vinci (NASA, 2016). He marvelled at how birds were able to fly in the sky while he was forcibly earthbound. For many years, he studied birds, focusing on how they flew and their wing structure. With all his knowledge, he set out to design one of the most complicated machines of his time and although his fabled "Ornithopter" was never constructed, his detailed blueprints laid down the framework for the development of fluid dynamics as a field. The earliest functional flying machines were produced between the 1800s and the 1850s (NASA, 2016; Gray, 2016). They were simple fixed-wing gliders created by George Cayley using the many prototypes method. The many prototype approach involves the development and testing of a design, observing what had worked, what had not, and then improving on the original design for the next prototype. This was the approach he had to take because the mathematics prior to this development suggested that flight was impossible. This was evident in 1749, when French mathematician Jean le Rond d'Alembert used Euler's equations for low-viscosity fluids, which assumed an incompressible and inviscid fluid with zero friction along the surface, to demonstrate that the high pressure in the front of the sphere would be balanced out by an equal high pressure at the rear of the sphere in Figure 1.

This meant that the air flowing around a sphere would not create any net drag. He was aware that was untrue in the physical world, but d'Alembert would never be able to explain this disconnect between his math and the world, stating that:

"It seems to me that the theory (potential flow), developed in all possible rigour, gives, at least in several cases, a strictly vanishing resistance, a singular paradox which I leave to future Geometers to elucidate" (Claes Johnson, 2012a).

This conundrum was later designated "d'Alembert's paradox", and split the field of fluid studies into the physical-world observation-based field of hydraulics, and the theoretical, math-based field of fluid mechanics. George Cayley's many prototypes method allowed for physical observation including ideal wing shapes, the best angle of attack, and other ways to increase lift. These observations were used to derive equations that allowed for calculations, but how exactly the small changes he made to the design of his prototype actually kept the plane in the air was still

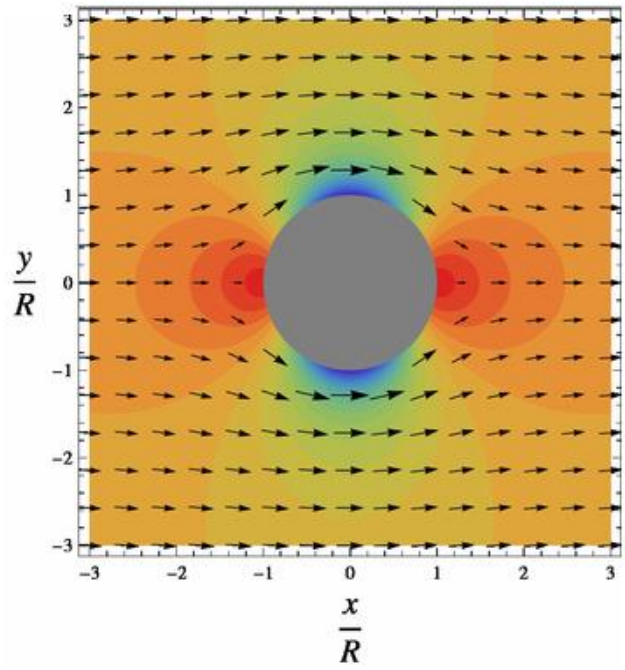


Figure 1: The potential flow solution according to the Navier-Stokes equation, with blue areas representing areas with low pressure while red areas represent areas of high pressure. Stokes and d'Alembert both found that the two areas of high pressure push against the sphere and cancel each other out resulting in no net drag, a conclusion which did not agree with observations made in the physical world (Claes Johnson, 2012a). (Image adapted from Incredio, 2009).

not yet understood. However, this method did pay off for George Cayley because after 50 years of prototypes he had finally designed a glider that was capable of flight in 1849 (Gray, 2016).

In 1822, the Navier-Stokes equations were developed by Claude-Louis Navier and George Gabriel Stokes. These equations were created to describe how a velocity vector field affects a fluid, or in terms of flight: how air flows around the wing. In 1851, Stokes, using these equations, modelled the flow of a fluid around the sphere in

Figure 1 (Incredio, 2009). He found that if he decreased the viscosity of the fluid, the drag on the sphere would also decrease, which was similar to what d'Alembert had found. This was more evidence supporting the idea that low-viscosity air around a wing would not produce any drag, which did not reflect reality. At this point, George Cayley had proven that flight was possible, while theoretical fluid mechanics had barely taken off.

PRANTL'S THEORY OF NO SLIP BOUNDARY

This disconnect continued to grow when the Wright brothers, Orville and Wilbur Wright, built a flying machine capable of self-powered flight in 1903 (NASA, 2016). It was a simple machine, but it was capable of sustaining itself in the air for 59 seconds and had travelled a total distance of 852 feet. This was a great accomplishment, which mathematics had determined to be theoretically impossible. This achievement brought new life into the field of fluid dynamics as scientists rushed to come up with a theory of how the Wright Brothers made their machine fly. One of the first individuals who attempted to rectify the disconnect between the mathematics of flight and physical flight was the German physicist Ludwig Prandtl. In 1904, he published his theory of flight (Hoffman and Johnson, 2008; 2009). He knew of d'Alembert's work and stated that his assumption of a slip-boundary was incorrect. He believed that at low viscosities, the surface and the molecules of the fluid would still interact and exert a friction force against the direction of fluid flow.

His theory of flight suggested that across the surface of a wing, a very thin layer of air would form, which he called the boundary layer. The part of the boundary layer that was immediately adjacent to the solid surface was called the no-slip boundary layer (Hoffman and Johnson, 2009; Claes Johnson, 2012b). He suggested that in this no-slip boundary layer the molecules of the fluid would collide and stick to the surface as it moved across it, meaning that the velocity of the molecules in this layer would move at the same velocity as the surface it was attached to. Meanwhile, for the rest of the boundary layer, the molecules of the fluid would also experience the effects of friction, but these effects would decrease to zero when the molecules were sufficiently far away from the surface of the wing, at the end of the boundary later. This assumption was based on the belief that the strength of adhesion between the solid boundary and the molecules were larger than the strength of cohesion between molecules. Using this assumption, he built a set of equations that could be used for this boundary condition, which was different than the general Navier-Stokes equations for the fluid outside of this boundary. Thus, Prandtl had shown that d'Alembert had not accounted for the formation of this boundary layer near the surface and that the drag was due to the change of momentum of the molecules in the boundary layer experiencing friction. Prandtl's resolution was also able to determine the source of turbulence as the collision of the different velocity molecules in the boundary layer. This disrupts the normally laminar boundary condition into a turbulent boundary condition once a threshold is overcome (Hoffman and Johnson, 2009).

KUTTA-ZHUKOVSY THEORY OF CIRCULATING AIRFLOW

Soon after Prandtl, Martin Wilhelm Kutta and Nikolai Zhukovsky separately published their theories of lift (Hoffman and Johnson, 2009). They thought that flights were due to a difference in pressures on the wing. They assumed that air would be trapped circulating around the wing going from over top the surface and rotating into opposite direction of movement beneath the bottom of the wing; this is depicted in Figure 2 (Claes Johnson, 2012b).

Their model found that the addition of this circulating air would cause the velocity of the molecules to increase in speed, in the direction of vortex movement, at the top of the wing, and slow down as it moves against the curl of the vortex beneath the wing. This would cause a difference in pressure favouring the upper surface of the wing, which would form a pressure gradient that would force the wing upwards (Hoffman and Johnson, 2009).

Exactly how the circulating air was formed was not clear, but it was thought to be the result of the sharp trailing edge guiding the motion of the air

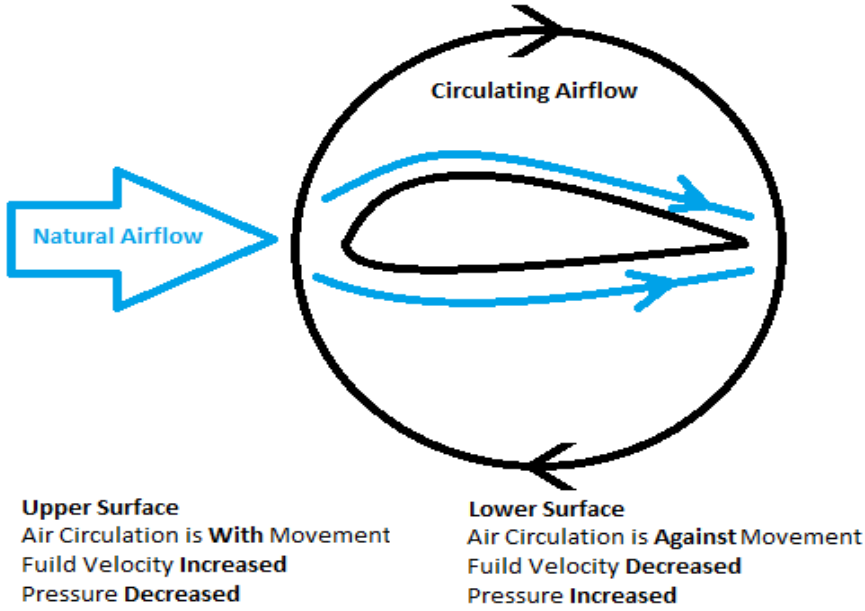


Figure 2: Kutta-Zhukovsy's circulation theory suggests that there was a circulating airflow which revolved around the wing. The molecules above the wing would move in the same direction as the circulating airflow and would result in a low pressure on the upper surface of the wing. The circulating airflow at the lower surface of the wing would act against the natural direction of airflow, which caused a retardation of velocity and would result in an increase in pressure. This pressure gradient between the upper and lower surface acts on the wing to create lift. Image adapted from (Hoffman and Johnson, 2009).

around the wing (Claes Johnson, 2012b). Though neither theory was confirmed, both Prandtl's theory of drag and Kutta-Zhykovsky's theory of lift would be used in calculations for 100 years and was sufficient to design aerodynamic wings and to further develop the field of fluid mechanics. For 100 years, this was the most widely accepted theory of flight and was found in all advanced fluid dynamic textbooks (Hoffman and Johnson, 2009; Claes Johnson, 2012b).

HOFFMAN AND CLAES' INSTABILITY SOLUTION

Before any steps can be made towards understanding the mathematics of flight, a basic understanding of vector calculus and fluid mechanics is required.

Vectors are simply a subset of tensors known as first-rank tensors. On a similar vein, scalars are called zeroth-rank tensors. Tensors are also composed of a magnitude and a direction. Similarly to how vectors can describe forces at a point, a tensor describes the forces acting on a surface. Taking a hypothetical solid shape, it can be assumed that the forces on it could be represented as the net vector force acting on the center of mass. However, in fluid dynamics, this interpretation is too limited because it does not take into account the movement of fluid flowing across a surface. This movement will cause a force on the surface that is parallel to the direction of the surface. Thus, tensors are better used to represent

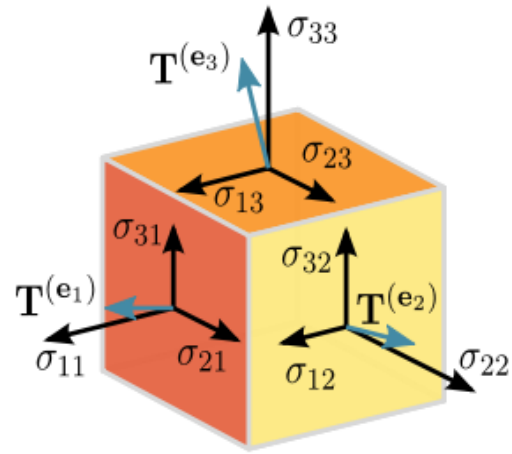


Figure 3: Each Tensor T is composed of 3 other vectors, σ, which show how the plane on each face are affected by some force. Image adapted from (Sanpaz, 2011)

the mathematics behind fluid motion. This also means that each direction of the space would have 3 vectors built into the tensor for the face orthogonal to that direction, resulting in a net total of 9 different vectors; this can clearly be shown in Figure 3.

VARIABLES

Hoffman and Johnson were able to solve d'Alembert's paradox using 5 simple equations. These equations involve the variables: u, which is a velocity vector composed of u_1,u_2,u_3, also written as $u = (u_1,u_2,u_3)$; pressure, p; viscous shear stress, σ; volume force, f; the net flow velocity, g; and β, which is the small skin friction coefficient

(Hoffman and Johnson, 2008; 2009; Claes Johnson, 2012c).

EQUATIONS

The following 5 equations are the ones which Johan Hoffman and Claes Johnson started with to proof their theory:

$$\dot{u} + (u \cdot \nabla)u + \nabla p - \nabla \cdot \sigma = f \qquad in\ \Omega \times I \qquad (1)$$

$$\nabla \cdot u = 0 \qquad in\ \Omega \times I \qquad (2)$$

$$u_n = g \qquad on\ \Gamma \times I \qquad (3)$$

$$\sigma_x = \beta u_x \qquad on\ \Gamma \times I \qquad (4)$$

$$u(\cdot,0) = u^0 \qquad in\ \Omega \qquad (5)$$

First, it is important to note the dimensions to which these equations apply. The first 2 equations involve a volume, Ω, in $\mathbb{R}3$ and a time interval, I, while the next 2 equations involve the boundary surface, Γ, that encompasses Ω and time interval I. Essentially, the first 2 equations describe properties of a volume and the next 2 equations describe the surface of that volume. The last equation deals with the volume Ω at time 0 because this is standard notation for setting an initial variable condition. Equation 1 is the Navier-Stokes equation solved for force, f. In this equation, $\dot{u}$ denotes the derivative of velocity with respect to time, $\frac{d}{dt}u(t)$. The second equation shows that the divergence velocity, $\nabla \cdot u$, of the fluid is 0. This is based on the assumption that the fluid is incompressible, one which all other major theories have made. Equation 3 states that the normal velocity, u_n, of the fluid at the boundary is equal to the net flow velocity g across the boundary. This quantity is zero for an airplane wing, because no air is flowing directly into or out of the wing. In the fourth line, we have that the tangential stress, σ_s, is equal to the tangential velocity, u_s, multiplied by skin friction coefficient, β (Hoffman and Johnson, 2008; 2009). Here we note that Prandtl had set the value of β to be large due to his assumption of a no-slip boundary, and stating that the tangential friction between the air and the wing was what caused the formation of drag. D'Alembert, Johan Hoffman, and Claes Johnson, on the other hand, assumed that there was no friction at the boundaries ($\beta = 0$). Solving for potential flow, Hoffman and Johnson got the same result as d'Alembert: a lack of drag $f = 0$ (Hoffman and Johnson, 2008; 2009; Claes Johnson, 2012c). However, what they did realize was that this result only existed in theoretical situations, and that in the real world, this solution was unstable due to the two points of instability in the equation (Hoffman and Johnson, 2008; 2009; Claes Johnson, 2012c). The first point being: $\sigma = \nabla \cdot (2\nu\varepsilon(u))$ and $\sigma_s = 2\nu\varepsilon(u)_s$, where ν is the viscosity of the fluid and $\varepsilon(u)$ is the usual velocity strain (Hoffman and Johnson, 2008; 2009). Although small, the effect of viscosity in these equations is what causes the potential solution to fail as a physical solution.

This instability can be visualized by considering the difference between two different yet similar states. First, let $(v, q, \tau) = (u - \bar{u}, p - \bar{p}, \sigma - \bar{\sigma})$ then the new equations appear in the form (Hoffman and Johnson, 2008; 2009):

$$\dot{v} + (u \cdot \nabla)v + (v \cdot \nabla)\bar{u} + \nabla q - \nabla \cdot \tau = f \qquad in\ \Omega \times I \qquad (6)$$

$$\nabla \cdot v = 0 \qquad in\ \Omega \times I \qquad (7)$$

$$u \cdot n = g - \bar{g} \qquad on\ \Gamma \times I \qquad (8)$$

$$\tau_x = 0 \qquad on\ \Gamma \times I \qquad (9)$$

$$v(\cdot,0) = u^0 - \bar{u}^0 \qquad in\ \Omega \qquad (10)$$

The key part of these equations is the $\nabla\bar{u}$, which is the divergence of $\bar{u}$. As shown by the incompressibility assumption in equations 2 and 7, $\nabla\bar{u}$ must also equal 0. As stated before, u is a velocity vector built with component functions: (u_1, u_2, u_3), thus $\bar{u}$ must also be built from similar component functions $(\bar{u}_1, \bar{u}_2, \bar{u}_3)$ then $\nabla\bar{u}$ is a 3 by 3 matrix in the form:

$$\nabla\bar{u} = \begin{vmatrix} \frac{\partial \bar{u}_1}{\partial x} & \frac{\partial \bar{u}_1}{\partial y} & \frac{\partial \bar{u}_1}{\partial z} \\ \frac{\partial \bar{u}_2}{\partial x} & \frac{\partial \bar{u}_2}{\partial y} & \frac{\partial \bar{u}_2}{\partial z} \\ \frac{\partial \bar{u}_3}{\partial x} & \frac{\partial \bar{u}_3}{\partial y} & \frac{\partial \bar{u}_3}{\partial z} \end{vmatrix}$$

When finding the trace of this matrix, it was found that it was equal to the equation for the divergence, which was 0. This results in the trace of this matrix also being equal to 0. Since the trace is also defined as the sum of the possible eigenvalues, or solutions to the matrix, there must be both positive and negative eigenvalues. This plays a major role in

finding the vorticity solution of this situation. By applying the cross divergence operator $\nabla \times$ to the Navier Stoke's equation, we can get the vorticity ω equation, where it is important to know that $\omega = \nabla \times u$. Since it is a cross product, the direction of the end vector depends on u, which means that ω is highly affected by the sign of the solution. Since both positive and negative solutions are possible it is easy to see how the potential solution is unstable.

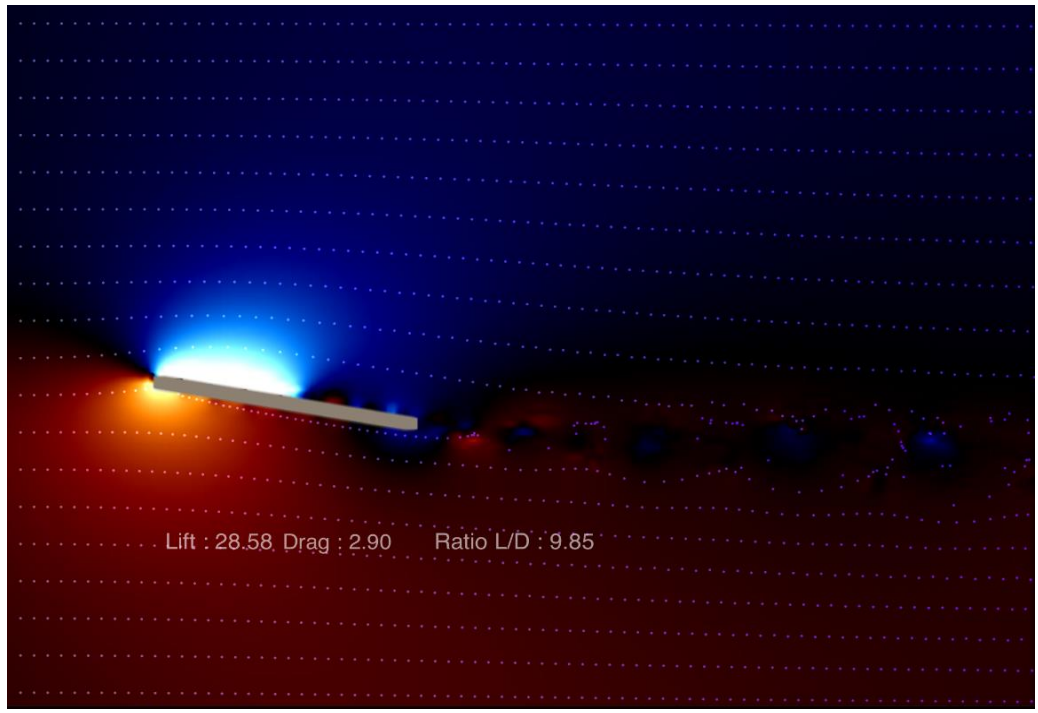


Figure 4: A simulation using the Wind Tunnel software showing that the intensity of the lift force (blue, low pressure) is about 9.85 times as big as the intensity of the drag force (red, high pressure). There is also no high pressure at the trailing edge of the wing, which was observed in Figure 1 (Algorizk, 2015).

Thus, in their paper "*The Mathematical Secret of Flight*", Hoffman and Johnson concluded that flight was only possible due to the production of a large lift to drag ratio, $\frac{L}{D}$ (see Figure 5). This ratio compares the amount of lift generated to the drag produced, therefore a ratio of 10 would mean that a plane is able to travel 10 metres for every 1 metre it descends (Hoffman and Johnson, 2008; 2009; Claes Johnson, 2012c). Their theory of flight was centered around 3 main ideas. These include the attachment of air to the leading edge of the wing, increasing the pressure against the motion of the wing. The lift created through the pressure drop at the top of the wing causes air to push upwards on the wing. Both of these are evident in d'Alembert's and Stokes' potential solutions; however the issue they had was that the potential solution also had an additional high pressure zone behind the solid surface. Hoffman and Claes were able to realize that the potential solution behind the airfoil was very unstable since it is caused by the collision of different flows of fluid, therefore in real life any small perturbations of the flow around the wing would cause the potential flow at the back of the wing to collapse into a different equation. They found that instead of potential flow at the trailing edge, a rotational flow would form instead due to the collision of molecules behind the trailing edge from the bottom and top of the wing. This created an oscillating velocity and the formation of alternating, counter-rotating pressure vortices (see Figure 6) (Hoffman and Johnson, 2009; Claes Johnson, 2012c).

These vortices would form with a low pressure inside that would negate the high pressure surrounding them. This instability mechanism is what decreases the pressure at the trailing edge of the wing and causes a net high pressure pushing against the wing, which is where the drag originates. The forward attachment of the flow is

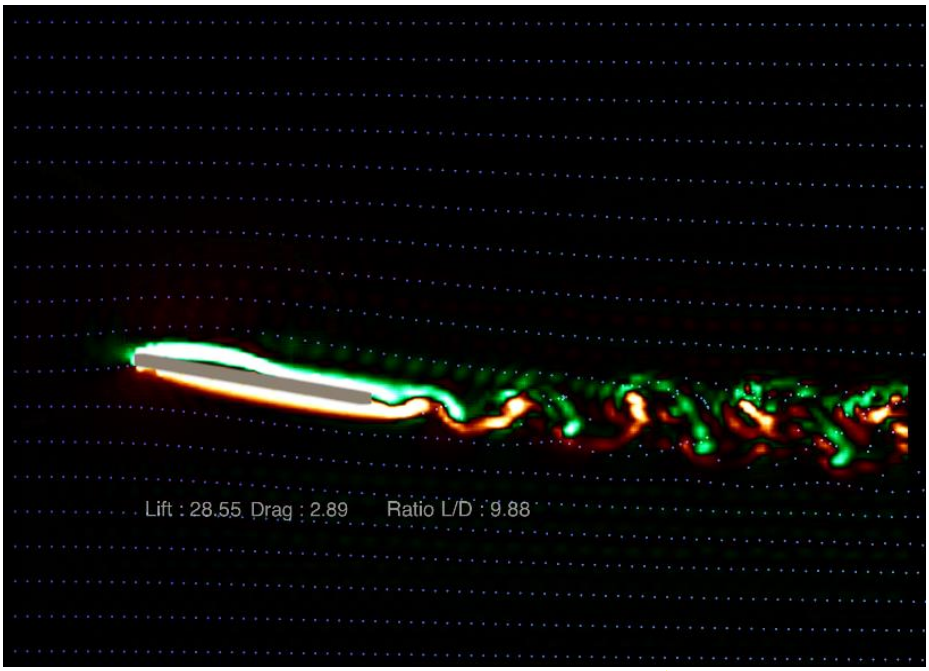


Figure 5: A simulation using the Wind Tunnel software showing that the alternating counter rotating pressure vortices at the trailing edge of the wing (Algorizk, 2015).

physical, due to the solid boundary and not from opposing flows, which results in much smaller perturbation growth than at rear separation, which is why vortices only form at the trailing edge (Hoffman and Johnson, 2008; 2009; Claes Johnson, 2012c).

CONCLUSION

Finally, after 250 years, a coherent resolution to d'Alembert's paradox has been found. After proving Prandtl and Kutta-Zhukovsy incorrect, Hoffman-Johnson reconciled the mathematics of field dynamics with observations in the physical world. They were able to determine that d'Alembert's potential was not stable due to the creation of both positive and negative solutions to the vorticity equation. This instability causes the formation of counter-rotating vortexes that counteract the high pressure at the trailing edge of the wing. This results in a net drag pushing against at the movement of the wing, a simple physical observation that has only now been given mathmatical support. Their solution has been tested in many computations and has correctly mimicked the peculiarities of real life. Those who are scared of flying may now breathe a sigh of relief as we can finally fully answer "What keeps planes up?". Thanks to this finding, airplane developers are take greater steps to improving the design of airplane wings and make air travel much more efficient.

ACKNOWLEDGEMENTS

I would like to thank George Dragomir for being my supervisor in this project as well as thank the Integrated Science Program for giving me this opportunity to do this research.

REFERENCES

Claes Johnson, 2012a. *D'Alembert's Paradox.* [video online] Available at: <https://www.youtube.com/watch?v=QPTjuS5G3dY > [Accessed 16 March 2016].

Claes Johnson, 2012b. *Incorrect or Trivial Theories of Flight.* [video online] Available at: <https://www.youtube.com/watch?v=LNLXdlXKvL8> [Accessed 16 March 2016].

Claes Johnson, 2012c. *New Theory of Flight.* [video online] Available at: <https://www.youtube.com/watch?v=t7e_6bkUFzE > [Accessed 16 March 2016].

Gray, C. 2016. *Flying Machines - Sir George Cayley.* [online] Flying Machines. Available at: <http://www.flyingmachines.org/cayl.html> [Accessed 16 March. 2016].

Hoffman, J. and Johnson, C. 2008. Resolution of d'Alembert's Paradox. *Journal of Mathematical Fluid Mechanics*, [online] 12(3), pp.321-334. Available at: <http://link.springer.com/article/10.1007/s00021-008-0290-1> [Accessed 15 Mar. 2016].

Hoffman, J. and Johnson, C. 2009. The Mathematical Secret of Flight. *Normat*, [online] 57(4), pp.1-25. Available at: <http://www.csc.kth.se/~cgjoh/flightnormat-5.pdf> [Accessed 16 Mar. 2016].

Incredio, 2009. *Flow Past a Cylinder. Pressure field and velocity vectors* [image online] Available at: <https://en.wikipedia.org/wiki/Potential_flow_around_a_circular_cylinder#/media/File:CylinderVelocityPressure.png> [Accessed 16 March 2016].

NASA 2016. *History of Flight.* [online] Available at: <https://www.grc.nasa.gov/www/k-12/UEET/StudentSite/historyofflight.html>[Accessed 12 Nov. 2016].

Sanpaz, 2011. *Components of the Cauchy stress tensor in Cartesian coordinates.* [image online] Available at: <https://en.wikipedia.org/wiki/Tensor#/media/File:Components_stress_tensor.svg> [Accessed 16 March 2016].

Algorizk 2015. *Wind Tunnel v1.3.2.* iPad [Accessed 16 March 2016].

Emerald Ash Borer Incidence and Infestation at McMaster Forest Teaching and Research Facility

Kyra Simone
Integrated Science Program, Class of 2017, McMaster University

SUMMARY

The emerald ash borer (EAB), *Agrilus planipennis*, poses great risk to Canada's ash trees (*Fraxinus* spp.); it threatens forested areas, urban shade trees, and manufacturing and shipping industries. EAB interferes with interactions of native species and have the potential to initiate ecosystem-wide cascades. This study assesses the current state of EAB infestation at McMaster Forest Teaching and Research Facility and attempts to elucidate relationships between EAB incidence and ecological factors. Ash trees were selected using stratified random sampling within pre-existing habitat classes. Selected trees were surveyed for evidence of EAB infestation. In addition, tree location, diameter at breast height (DBH), height of lowest exit holes, and visual assessments of tree health were recorded. Results do not indicate a clear effect of ecological land classification on EAB infestation across the McMaster Forest. Evidence of EAB activity is prevalent throughout the Forest and across all surveyed land classes: deciduous forest, deciduous woodland, mixed plantation, mixed forest, and deciduous shrub thicket. Data suggest that insect-foraging bird damage may be a useful indicator for future assessment of EAB infestation. The results of this study highlight the current state of EAB infestation in McMaster Forest, solidify some principles of the visual techniques used to assess EAB infestation, and provide insight into EAB distribution in relation to several ecological factors.

Received: 02/02/2016 **Accepted:** 03/25/2017 **Published:** 04/06/2017
URL: https://journals.mcmaster.ca/iScientist/article/view/1152/1005

Keywords: *Fraxinus, Agrilus*, invasive species, ecological land classifications, Canada, urban forest

INTRODUCTION

Ash trees have great ecological, economic, and aesthetic value in Canada. They are not only extremely widespread in forested areas, but also frequently planted as urban shade trees (Poland and McCullough, 2006). These trees increase property value, cool urban environments, take up storm water, and reduce airborne pollutants. Ash wood is also used in the manufacture and shipping of many products (Timms, et al., 2006). Unfortunately, the Asia-native emerald ash borer (EAB), *Agrilus planipennis*, has decimated millions of North America's ash trees since its introduction presumably around 2002 (Poland and McCullough, 2006).

Shipping infested wooden pallets or crates likely facilitated international migration of the flat-headed, wood-boring EAB. While this infestation was initially detected in Michigan and Ontario, EAB has since spread to 21 US states and two Canadian provinces (Xie, et al., 2004). It is considered the most costly biological invasion by a forest insect in North American history (Herms and McCullough, 2014). In addition to economic impacts, EAB infestation has immense and potentially cascading ecological effects. These

beetles produce canopy gaps, increase coarse woody debris, and alter delicate understory environments and nutrient cycling. Over 40 North American arthropod species feed on ash and are now at risk of local extinction as ash populations decline (Herms and McCullough, 2014).

The invasive ability of EAB is partially a result of stratified dispersal; by guerrilla strategy, the beetle invades new areas by flying to different sections of canopy, but also covers large distances via human transport of infested wood (Lopez, Acosta and Serrano, 1994). This allows small populations to be established at great distances from the initial infestation and accelerates radial spread (Liebhold and Tobin, 2008).

Initially, EAB infests the ash tree canopy. Lower regions are affected only later on, as tree health declines (Cappaert, et al., 2005). EAB activity is usually greatest in June and July, during which time adults feed for approximately 1 week before mating. Within their short 3–6-week lifespan, females lay between 50 and 90 cream-coloured eggs in bark crevices (Poland and McCullough, 2006). Within a few days, eggs darken to reddish-brown (Cappaert, et al., 2005). Larvae emerge from eggs within 2 weeks and begin to feed on phloem. As the larvae feed, they carve out winding galleries that become filled with sawdust-like frass. Eventually, this tunnelling becomes so extensive that it disrupts tree nutrient transport mechanisms (Poland and McCullough, 2006). Under such stress, trees may develop epicormic shoots: new branch or leaf growths below the point of infestation. Infested ash trees are eventually unable to circulate water and nutrients, and usually die within 1–3 years (Poland and McCullough, 2006).

EAB overwinters within ash trees as either pre-pupae or young larvae; the latter require a second year of development before emergence as adults. In spring, the adult EAB chews through bark to emerge from the tree. This occurs once temperatures stabilize above approximately 10°C. They leave behind D-shaped exit holes, about 3–4 mm in width (Poland and McCullough, 2006). These are easily distinguishable from exit holes of native wood-borers (**Figure 1**), such as the ash bark beetle, red-headed ash borer, and banded ash borer. Native borers leave behind much rounder holes which tend to be substantially larger or smaller than those caused by EAB (Burns and Honkala, 1990; Fuester, et al., 2007; Stepanek, 2014). Native borers typically only infest stressed trees and thus do not cause anywhere near as much devastation to ash populations. Similarly, in Asia, EAB attacks unhealthy trees. In North America, however, EAB infests even healthy ash trees (Herms and McCullough, 2014).

The purpose of this study is to determine the current state of EAB infestation at McMaster Forest Teaching and Research Facility in Hamilton, Ontario, and to elucidate any potential relationship between EAB distribution and ecological factors. McMaster University purchased the 48-hectare Forest in 1964. It consists of grassland, forest, and wetland regions, and is used as a study site for various biodiversity studies and restoration projects. More specifically, the Forest is divided into 14 different regions, each defined by specific ecological factors (**Figure 2**). Ash trees are very prevalent in the Forest; distributed sparsely in the front, prairie region, and closely aggregated in the forested, rear area. This, alongside the sheer number of ecosystems represented, makes the Forest an extremely useful study site. Its biodiversity is analogous to that of Southern Ontario, and allows this study to make inferences about larger-scale patterns.

Figure 1: Comparison of D-shaped EAB exit holes (left), round ash bark beetle holes (centre), and round banded ash borer holes (right) (de Groot, et al., 2006; Fuester, et al., 2007; Stepanek, 2014).

METHODS

Data collection took place between mid-January and mid-March, 2016, and included measurements of diameter at breast height (DBH), tree location, height of lowest EAB exit hole or larval gallery above the ground, and visual observations of tree health. Visual observations included dead branches, woodpecker or squirrel damage, and bark cracks or

peeling. Standing, dead ash trees were also assessed.

The entire McMaster Forest site is divided into 100-metre grid squares, while the back, forested area is further sub-divided into 20-by-20 metre quadrats. An ongoing census of the Forest has thus far catalogued over 16,000 trees, including data concerning species (in accordance with the Ministry of Natural Resources' vascular plant species list), DBH, and location (x,y coordinates). This study obtained data from the ecological land classifications represented by the rear quadrats. Stratified random sampling was used to select quadrats from each of the land classes in which ash trees grow. Other researchers have conducted EAB surveys using similarly random sampling, such as random transects across the area of study (Smitley, et al., 2008). In this study, data was collected from all trees in 5 random quadrats from each ecological class.

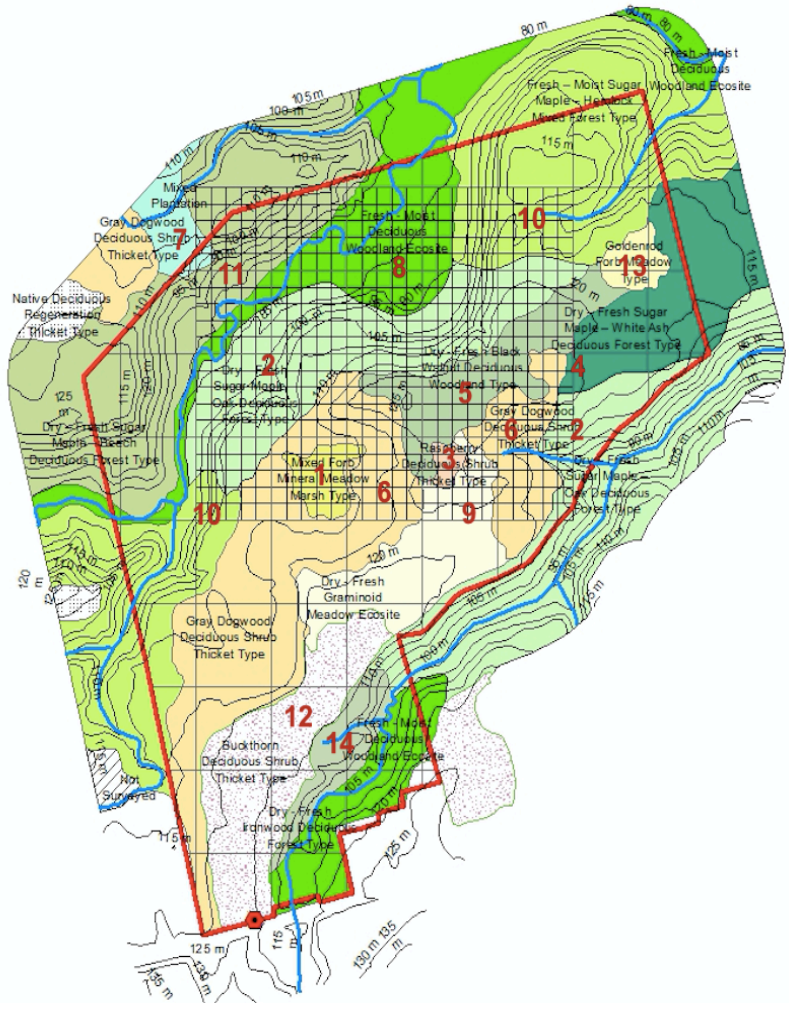


Figure 2: Ecological regions of McMaster Forest. 1) Mixed forb mineral meadow marsh, 2) Dry fresh sugar maple oak deciduous forest, 3) Raspberry deciduous shrub thicket, 4) Dry fresh sugar maple white ash deciduous forest, 5) Dry fresh black walnut deciduous woodland type, 6) Gray dogwood deciduous shrub thicket, 7) Mixed plantation, 8) Fresh moist deciduous woodland ecosite, 9) Dry fresh graminoid meadow ecosite, 10) Fresh moist sugar maple hemlock mixed forest, 11) Dry fresh sugar maple beech deciduous forest, 12) Buckthorn deciduous shrub thicket, 13) Goldenrod forb meadow, 14) Dry fresh ironwood deciduous forest.

SYMPTOMS OF EAB INFESTATION

Visual assessment techniques largely detect only later-stage EAB infestation (Ryall, et al., 2011). Dendrochronological data has shown that EAB in Michigan remained undetected for a decade, likely because native borers produce similar symptoms in affected trees (Cappaert, 2005). This study acknowledged such limitations, but continued to utilize non-destructive, visual observation techniques to provide a preliminary assessment of EAB infestation in the McMaster Forest.

More specifically, this assessment involved close binocular examination of ash trees, which has been effective in previous studies (Smitley, et al., 2008). EAB is also more likely to attack trees infested by other beetles, likely because it is attracted by stress-induced pheromone production (Cappaert, 2005). Thus, evidence of other beetle activity, including round exit holes, were considered to be indicative of tree vulnerability to infestation. Similarly, while woodpecker damage or bark peeling by squirrels could be due to feeding on many boring insects, these patterns could also be associated with EAB infestation (de Groot, et al., 2006). So, any observed drill-like woodpecker marks or ragged bark were noted.

As infestation occurs first in the canopy and progresses downward, the height of the lowest exit hole or larval gallery above the ground was recorded for each infested tree. This provided an approximate measure of the stage of infestation. Regions below cracked bark were closely observed for larval galleries. EAB seldom infests trees less than a certain diameter. This is likely because it requires substantial bark thickness to provide nutrition and protection from desiccation and temperature extremes (Timms, et al., 2006). Thus, ash trees of DBH 5 cm or greater were surveyed in each of the randomly selected quadrats.

Table 1: Proportion of ash trees infested in each land class.

Land Class	Proportion of Ash Infested
Mixed plantation	0.667
Fresh moist sugar maple hemlock mixed forest	0.619
Dry fresh sugar maple white ash deciduous forest	0.571
Dry fresh sugar maple oak deciduous forest	0.452
Fresh moist deciduous woodland ecosite	0.419
Dry fresh sugar maple beech deciduous forest	0.4
Dry fresh black walnut deciduous woodland type	0.333
Gray dogwood deciduous shrub thicket	0.031

Censused regions of the McMaster Forest already associated species names with tag numbers; this information guided data collection. Some ash trees were missing tags, and had to successfully be distinguished from other species. In these cases, ash were distinguished from other tree species by their opposite branching, "braided" bark, highly veined leaf scars, stout twigs, and "scaly" buds (Lyons, et al., 2007). As data collection took place during the winter, tree identification was carried out without observation of leaf shape.

RESULTS

While this preliminary survey did not include all land classes in the McMaster Forest, it provided a comprehensive assessment of the state of infestation at the research site. Evidence of EAB infestation was found in all 8 surveyed land classes, in the proportions shown in **Table 1**.

Initially, observation of EAB activity at the front of the Forest prompted a wider investigation of the extent of infestation. This resultant survey found that EAB infestation was not limited to the front region, and had actually spread to affect trees across the entire site (**Figure 3**). Infested ash trees were observed to often also display extensive woodpecker damage; almost 83% of infested trees also showed considerable bark mutilation by woodpeckers. In future surveys, this could serve as a useful indicator for potential EAB infestation.

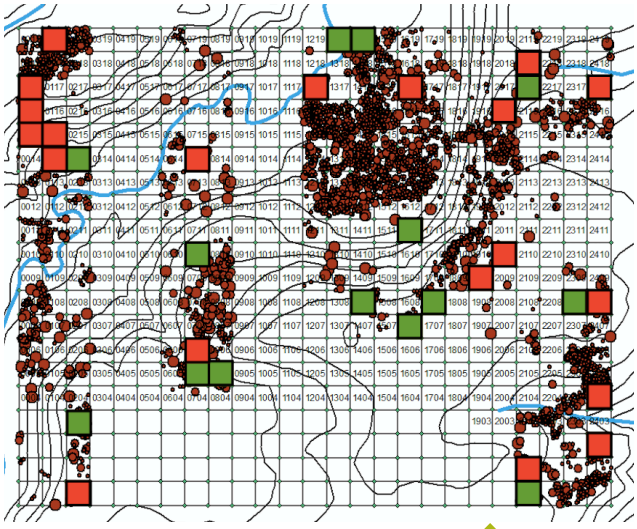


Figure 3: Red circles on map indicate censused ash trees (radii proportional to measured DBH). Quadrats indicated by coloured squares were selected randomly and surveyed. Red quadrats contained trees that exhibited signs of EAB infestation, while green quadrats did not.

A logistic regression was performed in order to determine any potential relationship between land class and EAB incidence. While a significant effect of land class on infestation was observed ($p = 0.00266$), a post-hoc Tukey's test revealed that only one land class was significantly less infested than any others (**Figure 4**). One land class, "gray dogwood deciduous shrub thicket", remained relatively uninfested, while all other surveyed quadrats displayed 33% or greater infestation rates. These results, while preliminary, suggest that ecological factors do not play a great role in EAB infestation pattern. However, it is possible that the gray dogwood thicket region has several characteristics, including lower tree density, that make its ash trees less susceptible to infestation.

Regional characteristics did not seem to greatly affect EAB distribution in the Forest. However, some properties of ash trees appeared to affect EAB incidence. Previous studies have not found infestation to be limited to trees within a certain DBH range (Cappaert, et al., 2005). While this survey did not find evidence directly to the contrary, it appears that EAB tended to infest ash trees with mature, braided bark.

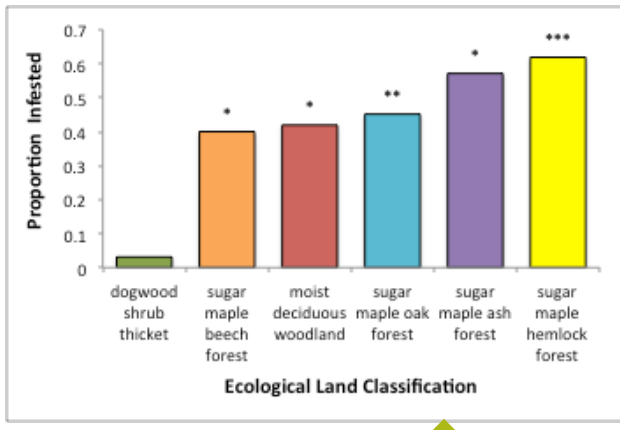


Figure 4: Significant differences between proportions of infested trees in various land classes (p-value < 0.001 represented by "***", p < 0.01 represented by "**", p < 0.05 represented by "*"). The only significant differences were found between gray dogwood thicket and several other highly infested land classes, suggesting little association between ecological factors and EAB infestation patterns.

Younger, smaller ash with smooth bark tended not to exhibit as many signs of infestation. This supports the theory that EAB prefer bark of a certain thickness for its protective capabilities (Timms, et al., 2006). Furthermore, older, larger ash with extremely ridged bark tended to also exhibit less signs of infestation. Future studies could investigate the potential for such lower and upper DBH limits, within which EAB infestation might be most likely.

DISCUSSION

FUTURE DIRECTIONS

MCMASTER FOREST: Ideally, a more in-depth survey of the McMaster Forest would be carried out to better estimate the effects of ecological factors on EAB infestation. The limited sample size of this study makes accurate prediction of such relationships difficult, especially over larger spatial scales. A larger survey would help to elucidate any relationship, or the lack of any relationship, between EAB infestation and ecological factors, both of which would be extremely informative. This information could then be scaled up to larger geographical scales, where it could influence regional policy development.

As this preliminary survey indicates, EAB infestation is widespread across the McMaster Forest. Ash trees make up a significant fraction of the species in this study site, and their rapid decline will undoubtedly cause significant ecosystem impacts. Furthermore, infestation has the potential to spread from the Forest site as adult beetles emerge and fly to new host trees. McMaster faces a difficult decision in the near future, both ethically and financially. As a preliminary measure, it may be best to place the Forest site under local quarantine in the near future. Site users, including hikers, should be informed and urged not to transport wood into or out of the site.

Subsequently, more drastic measures may have to be employed, including pheromone or insecticide treatments, or even large-scale ash tree removal. Such measures should be implemented prior to adult emergence in late spring in order to prevent further spread.

Several insecticides effective against EAB are publicly available, but cannot protect trees already in late stages of infestation. These are most commonly applied as trunk or soil injections, or as spray directly onto the tree (Herms, et al., 2014). These methods are both labour-intensive and costly, as they require precise application to each and every ash tree. Furthermore, these compounds are toxic and can spread through soil and groundwater to affect flora and fauna in much larger areas (Hahn, Herms and McCullough, 2011). Ash bark volatiles have also been shown to effectively lure EAB into traps (Crook, et al., 2008). While this research is preliminary, perhaps in the future, such traps could be used to combat EAB infestation while also having a lesser impact on the surrounding ecosystem, as they would employ compounds endogenous to the ecosystem. As a last resort, the University could endeavour to physically remove all ash trees with mature, braided bark, or above a certain DBH. The wood material would then have to be burned on-site, to prevent any further spread of infestation.

Southern Ontario: On a larger scale, it would be prudent to implement pheromone traps across Southern Ontario to detect early stages of infestation (Abell, et al., 2015). Perhaps more urgently, these traps should be implemented in dense, urban ash monocultures, where EAB has the potential to wipe out many trees in a concentrated area.

Should future surveys indicate that certain ecological factors impact patterns of EAB infestation, these principles could be used to mimic such ecosystem characteristics and enhance resilience. However, until such results come to light, preventative measures must be implemented across all ecosystem types, with a focus on urban forests.

Further investigation of trunk injection efficacy against EAB infestations of varying severity would allow for more effective use of treatments such as TreeAzin®. Preliminary studies suggest that TreeAzin would be comparably effective in ash trees regardless of infestation severity, should greatly infested trees be treated more frequently (Herms et al., 2014).

Findings could then inform the introduction of an insecticide efficacy term into models of human-forest interactions and EAB spread. Such models demonstrate the importance of considering human involvement in stratified invasive species dispersal (*e.g.,* Ali, et al., 2015; Barlow, et al., 2014). Ideally, quantitative field results of an insecticide efficacy study, combined with conceptual representation of insecticide interference with EAB's stratified dispersal, would lead to implementation of more effective control strategies, especially in urban areas with many at-risk ash trees.

CONCLUSION

The emerald ash borer continues to pose great risks to biodiversity, industry, and property values in North America. The McMaster Forest site, with its diverse ecological composition and abundance of ash trees, provides a unique study environment for the extent and effects of this pest. This survey sheds light on EAB incidence in the McMaster Forest, as well as potential patterns for its distribution in Southern Ontario, through the observation of analogous ecosystems.

Measures should be implemented prior to late-spring emergence of adult beetles so as to prevent further spread of infestation and preserve remaining ash populations. These may include quarantines, insecticide application, or even large-scale ash tree removal.

ACKNOWLEDGEMENTS

Many thanks to Dr. Chad Harvey for his encouragement and guidance throughout this project. Geneviève van Wersch also aided greatly in the organization and communication of project expectations. Cross-referencing the ongoing tree census was immensely helpful in the completion of this survey. The supervision of Dr. Susan Dudley and Master's candidate Sophia Muñoz was greatly appreciated. Fieldwork could not have taken place without the help of Jordan Barker and Ann Marie Simone.

REFERENCES

Abell, K., Poland, T., Cossé, A. and Bauer, L., 2015. *Trapping Techniques for Emerald Ash Borer and its Introduced Parasitoids*. In: R. Van Driesche and R.C. Reardon, eds., Biology and Control of Emerald Ash Borer. Morgantown: Forest Health Technology Enterprise Team, pp.113–127.

Ali, Q., Bauch, C.T. and Anand, M., 2015. Coupled Human-Environment Dynamics of Forest Pest Spread and Control in a Multi-Patch, Stochastic Setting D. Doucet, ed. *PLoS ONE*, 10(10), pp.1–21.

Barlow, L.-A., Cecile, J., Bauch, C.T. and Anand, M., 2014. Modelling Interactions between Forest Pest Invasions and Human Decisions Regarding Firewood Transport Restrictions R. Planque, ed. *PLoS ONE*, 9(4), pp.1–12.

Burns, R.M. and Honkala, B.H., 1990. *Silvics of North America: Volume 2, Hardwoods*. 2nd ed. Washington, DC: United States Department of Agriculture.

Cappaert, D., McCullough, D.G., Poland, T.M. and Siegert, N.W., 2005. Emerald ash borer in North America: a research and regulatory challenge. *American Entomologist*, 51(3), pp.152–165.

Crook, D.J., Khrimian, A., Francese, J.A., Fraser, I., Poland, T.M., Sawyer, A.J. and Mastro, V.C., 2008. Development of a Host-Based Semiochemical Lure for Trapping Emerald Ash Borer Agrilus planipennis (Coleoptera: Buprestidae). *Environmental Entomology*, 37(2), pp.356–365.

de Groot, P., Biggs, W.D., Lyons, D.B., Scarr, T., Czerwinski, E., Evans, H.J., Ingram, W. and Marchant, K., 2006. *A Visual Guide to Detecting Emerald Ash Borer Damage*. Canadian Forestry Service: Sault Ste. Marie.

Fuester, R.W., Taylor, P.B. and Wildonger, J.A., 2007. *Insects Contributing to Ash Mortality in Eastern Pennsylvania*. University of Delaware: Newark.

Hahn, J., Herms, D.A. and McCullough, D.G., 2011. *Frequently Asked Questions Regarding Potential Side Effects of Systemic Insecticides Used To Control Emerald Ash Borer*. [online] East Lansing. Available at: <http://www.emeraldashborer.info/documents/Potential_Side_Effects_of_EAB_Insecticides_FAQ.pdf>.

Herms, D.A. and McCullough, D.G., 2014. Emerald Ash Borer Invasion of North America: History, Biology, Ecology, Impacts, and Management. *Annu. Rev. Entomol.*, 59, pp.13–30.

Herms, D.A., McCullough, D.G., Smitley, D.R., Sadof, C.S. and Cranshaw, W., 2014. *Insecticide Options for Protecting Ash Trees from Emerald Ash Borer*. 2nd ed. Urbana–Champaign: North Central IPM Center

Liebhold, A.M. and Tobin, P.C., 2008. Population Ecology of Insect Invasions and Their Management. *Annu. Rev. Entomol*, 53, pp.387–408

Lopez, F., Acosta, F.J. and Serrano, J.M., 1994. Guerilla vs. Phalanx Strategies of Resource Capture: Growth and Structural Plasticity in the Trunk Trail System of the Harvester Ant Messor barbarus. Source: *Journal of Animal Ecology*, 63(63), pp.127–138.

Lyons, D.B., Caister, C., de Groot, P., Hamilton, B., Marchant, K., Scarr, T. and Turgeon, J., 2007. *Survey Guide for Detection of Emerald Ash Borer*. Natural Resources Canada, Canadian Forest Service, Canadian Food Inspection Agency: Ottawa.

Poland, T.M. and McCullough, D.G., 2006. Emerald ash borer: Invasion of the urban forest and the threat to North America's ash resource. *Journal of Forestry*, 104(3), pp.118–124.

Ryall, K.L., Fidgen, J.G. and Turgeon, J.J., 2011. Detectability of the Emerald Ash Borer (Coleoptera: Buprestidae) in Asymptomatic Urban Trees by using Branch Samples. *Environmental Entomology*, 40(3), pp.679–688.

Smitley, D., Davis, T. and Rebek, E., 2008. Progression of ash canopy thinning and dieback outward from the initial infestation of emerald ash borer (Coleoptera: Buprestidae) in southeastern Michigan. *Journal of Economic Entomology*, 101(5), pp.1643–1650.

Stepanek, L., 2014. *Decline in Ash Trees: Borers and Bark Beetles An Identification Guide*. Nebraska Forest Service: Lincoln.

Timms, L.L., Smith, S.M. and Groot, P. de, 2006. Patterns in the within-tree distribution of the emerald ash borer Agrilus planipennis (Fairmaire) in young, green-ash plantations of south-western Ontario, Canada. *Agricultural and Forest Entomology*, 8, pp.313–321.

Xia, W., Reardon, D. Wu, Y. and JiangHua, S., 2004. Emerald ash borer, Agrilus planipennis Fairmaire (Coleoptera: Buprestidae), in China: a review and distribution survey. *Acta Entomologica Sinica*, 47(5), pp.679–685.

Current Understanding of Europa and Potential in Upcoming Exploration

JAMES LAI
Michael G. DeGroote School of Medicine, Class of 2019, McMaster University

SUMMARY

Europa, one of Jupiter's moons, has been an object of interest to many scientists, including astrobiologists, for some time. In particular, Europa has drawn interest due to evidence suggesting the presence of a subsurface liquid water ocean, which could potentially support life. Only recently, however, has concrete planning begun for exploration of this celestial body. These missions are described, followed by an overview of selected areas of research on Europa relevant to astrobiologists, along with discussion of planned investigations for upcoming missions to explore Europa. Finally, based on these potential benefits to scientific knowledge, it is argued that exploration of icy moons such as Europa should remain a priority as they provide opportunities for the study of astrobiology that cannot be offered by the current focus of most study, Mars.

Received: 02/27/2017 **Accepted:** 03/19/2017 **Published:** 07/31/2017
URL: https://journals.mcmaster.ca/iScientist/article/view/1447

Keywords: Europa, astrobiology, space exploration, Jupiter, Mars

Europa (Figure 1) is Jupiter's sixth-closest moon and was first discovered by Galileo Galilei in 1610. Visually, Europa is characterized as a smooth-looking satellite devoid of large-scale features such as craters and mountains, but covered with many reddish streaks known as lineae (Alexander, et al., 2009).

By the late 1900s, mounting evidence suggested that various surface features of Europa were the result of activity of its water-ice crust, similar to plate tectonic activity on Earth.

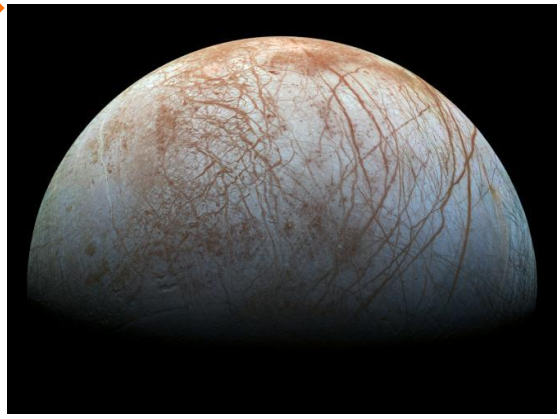


Figure 1: Europa, a moon of Jupiter, is known for its icy surface with reddish streaks known as lineae; this enhanced colour image was taken by NASA's Galileo mission (NASA, JPL-Caltech, SETI Institute, 2014).

Furthermore, this ice crust activity pointed to the presence of a layer of liquid water beneath the surface (Alexander, et al., 2009), drawing interest since liquid water is required for Earth-like life. With these advances occurring concurrently with the growth of astrobiology as a scientific discipline (see for example the overview in Lai, 2015), the potential for life in such an environment quickly caught the attention of astrobiologists.

Despite various proposed missions throughout the years, no concrete plans for a mission to Europa have been produced until recently. In light of these upcoming missions, an overview of currently planned missions will be given, followed by a summary of several topics of relevance to the astrobiological study of Europa, along with potential for scientific advancement with the upcoming missions. Finally, based on these benefits, it is argued that icy moon exploration should continue alongside exploration of Mars, since each environment provides very different scientific opportunities.

UPCOMING MISSIONS

One proposed missions to Europa is the European Space Agency's Jupiter Icy Moons Explorer. This mission would arrive in the Jovian system around 2030 and perform two flybys of Europa en route to studying Jupiter's moon Ganymede in greater depth (Grasset, et al., 2013).

The National Aeronautics and Space Administration (NASA) also has a planned mission to Europa: the Europa Multiple-Flyby Mission. This mission would arrive in the 2020s and conduct 45 flybys of Europa (NASA, 2016).

CURRENT UNDERSTANDING OF EUROPA

ICE CRUST THICKNESS

The thickness of Europa's ice crust remains a subject of debate. Estimates have ranged from a few hundred meters to several tens of kilometers (Hoolst, et al., 2008). Attempts to estimate the thickness of the ice crust have approached the problem by attempting to model various Europan phenomena, including tidal dissipation and surface features such as cycloid cracks and craters.

Tidal dissipation methods take into account the heat lost by Europa's crust and the heat generated by tidal forces on Europa due to Jupiter. Hussmann, Spohn, and Wieczerkowski (2002) used to method to calculate an ice crust layer that is a few tens of kilometers in thickness. More recently, Quick and Marsh (2015) used this method to estimate a thickness of 28 km.

The presence of cycloid cracks, which form pairs of ridges, has also been used as a method for estimating the thickness of Europa's crust. It has been suggested that in order for these features to form, Europa's crust thickness must be sufficiently thin that the cracks can reach a layer of liquid water. Based on calculations by made by Hoppa, et al. (1999), this means that the ice thickness must not exceed a few kilometers at most.

Crater depth methods make use of measurements of crater features on Europa's surface and attempt to model these to determine what crustal thickness would be consistent with the observed craters. Studies using this method have produced mixed results, but generally agree that the crustal thickness exceeds 4 km. In the work of Bray, et al. (2014), which uses this technique along with a hydrocode simulation, Europa's crustal thickness is estimated to be 7 km.

Given the variation in calculated ice thickness using different methods, it is evident that an upcoming mission could provide valuable direct measurements of the ice thickness. Both planned missions incorporate ice-penetrating radar as part of their instrument packages. Heggy, et al., (2017) predict that it should be possible to examine shallow subsurface features such as cracks with radar specifications like those proposed for these missions. Furthermore, with a possible penetration depth ranging from 1 to 18 km, it may be possible to characterize the deep ocean, and thus, the ice thickness.

PRESENCE OF SUBSURFACE OCEAN

Beneath the icy crust, it is proposed that Europa's subsurface contains a body of liquid water, which could potentially harbour life. The heat required to keep water in a liquid state in such conditions is proposed to be produced by tidal forces acting on Europa from Jupiter, along with other sources such as radioactive decay occurring at the putative seafloor (Melosh, et al., 2004). The composition of such an ocean is a topic of interest to researchers, with implications for both physical features observed on the surface, and for the conditions within which any potential organisms on Europa would exist (Marion, et al., 2004; Thomas, et al., 2017). With the upcoming missions, the characteristics of this ocean could be further elucidated. Such study can be done with instruments such as magnetometers, which could allow characterization of the ocean's salinity and depth (Raymond, et al., 2015; Lunine, 2017).

PLUMES AND POTENTIAL SAMPLING

While the presence of plumes that might vent the contents of the subsurface had been previously proposed, it was not until 2014 that evidence of their existence was found, in the form of UV emission detected by the Hubble Space Telescope (Roth, et al., 2014).

Sampling these plumes could allow direct study of a putative subsurface ocean; even more

intriguingly, such a sample could contain biomolecules, or even lifeforms. However, the potential for finding lifeforms is quite low. Lorenz (2016) created a simple model of such a situation, assuming sampling from a 2 km altitude, a supposed lower limit of feasibility. Based on this model, it would only be possible to detect a few hundred cells if Europa's waters were analogous to Earth's most life-rich waters; on the other hand, if Europa's waters were more like those of Lake Vostok, which is proposed to be a better analogue of Europan conditions, fewer than one cell would be statistically likely to be detected.

Upcoming missions have within their scope of interest the study of these plumes. These missions will characterize plumes with ultraviolet spectrometry; furthermore, any plume material that may be encountered could be examined using mass spectrometry (Lunine, 2017).

DISCUSSION AND CONCLUSION

Europa provides many exciting opportunities to further the understanding of life in the Universe. In particular, Europa's potential subsurface liquid water ocean could be a habitat for life, or provide clues as to conditions required for life to arise. Such an environment is not present on the target of greatest current astrobiological study: Mars. While evidence now suggests that Mars possessed liquid water oceans in the past (De Blasio, 2014, and references within), and may currently have some subsurface water ice that may melt (Stillman, Michaels, and Grimm, 2017), these conditions are very different from the water-rich environment that present-day Europa is theorized to possess.

In addition to providing an environment that might be more likely to currently harbour life than Mars, Europa, along with other icy moons such as Ganymede and Enceladus, also provides unique opportunities to study processes that are not present on Mars, but could affect the development of life, both here and elsewhere in the universe. Understanding of tidally-heated oceans and water ice crust dynamics could prove useful when applied to extrasolar systems, and the development of technologies required to detect biomolecules and signs of life could also further astrobiological research. Ultimately, exploration of icy moons like Europa should continue alongside the exploration of current targets like Mars, in order to broaden the perspectives available to astrobiologists in understanding the origins and distribution of life in the Universe.

REFERENCES

Alexander, C., Carlson, R., Consolmagno, G., Greeley, R., and Morrison, D., 2009. The exploration history of Europa. In: R. T. Pappalardo, W. B. McKinnon, K. Khurana, eds. 2009. Europa. Tucson: University of Arizona Press. pp.3-26.

De Blasio, F. V., 2014. Possible erosion marks of bottom oceanic currents in the northern lowlands of Mars. Planetary and Space Science, 93, pp.10-21.

Bray, V. J., Collins, G. S., Morgan, J. V., Melosh, J., and Schenk, P. M., 2014. Hydrocode simulation of Ganymede and Europa cratering trends – how thick is Europa's crust? Icarus, 231, pp.394-406.

Grasset, O., Dougherty, M. K., Coustenis, A., Bunce, E. J., Erd, C., Titov, D., Blanc, M., Coates, A., Drossart, P., Fletcher, L. N., Hussmann, H., Jaumann, R., Krupp, N., Lebreton, J.-P., Prieto-Ballesteros, O., Tortora, P., Tosi, F., and Van Hoolst, T. 2013. JUpiter ICy moons Explorer (JUICE): an ESA mission to orbit Ganymede and to characterise the Jupiter system. Planetary and Space Science, 78, pp.1-21.

Heggy, E., Scabbia, G., Bruzzone, L., Pappalardo, R. T., 2017. Radar probing of Jovian icy moons: understanding subsurface water and structure detectability in the JUICE and Europa missions. Icarus, 285, pp.237-251.

Hoppa, G. V., Tufts, R., Greenberg, R., and Geissler, P. E., 1999. Formation of cycloidal features on Europa. Science, 285(5435), pp.1899-1902.

Hussmann, H., Spohn, T., and Wierczerkowski, K., 2002. Thermal equilibrium states of Europa's ice shell: implications for internal ocean thickness and surface heat flow. Icarus, 156, pp.143-151.

Lai, J., 2015. Astrobiology: development of a science. In: C. Eyles, S. Symons, eds., 2015. History of the Earth Volume V. Hamilton: McMaster University. pp.38-41.

Lorenz, R. D., 2016. Europa ocean sampling by plume flythrough: astrobiological expectations. Icarus, 267, pp.217-219.

Lunine, J. I., 2017. Ocean worlds exploration. Acta Astronautica, 131, pp.123-130.

Marion, G. M., Fritsen, C. H., Eicken, H., and Payne, M. C., 2004. The search for life on Europa: limiting environmental factors, potential habitats, and Earth analogues. Astrobiology, 3(4), pp.785-811.

Melosh, H. J., Ekholm, A. G., Showman, A. P., and Lorenz, R. D., 2004. The temperature of Europa's subsurface water ocean. Icarus, 168(2), pp.498-502

NASA, 2016. Europa overview. [online] Available at: <https://www.nasa.gov/europa/overview/index.html> [Accessed 25 February 2017].

NASA, JPL-Caltech, and SETI Institute, 2014. PIA19048: Europa's stunning surface. [image online] Available at: <http://photojournal.jpl.nasa.gov/jpeg/PIA19048.jpg > [Accessed 21 February 2017].

Quick, L. C. and Marsh, B. D., 2015. Constraining the thickness of Europa's water–ice shell: insights from tidal dissipation and conductive cooling. Icarus, 253, pp.16-24.

Raymond, C. A., Jia, X., Joy, S. P. Khurana, K. K., Murphy, N., Russell, C. T., Strangeway, R. J., and Weiss, B. P., 2015. Interior Characterization of Europa using Magnetometry (ICEMAG): probing the Europan ocean and exosphere. American Geophysical Union, Fall Meeting 2015. Abstract only. Available through: SAO/NASA ADS Astronomy Abstract Service <http://adsabs.harvard.edu/abs/2015AGUFM.P13E..08R> [Accessed 26 February 2017].

Roth, L., Saur, J., Retherford, K. D., Strobel, D. F., Feldman, P. D., McGrath, M. A., and Nimmo, F., 2014. Transient water vapor at Europa's south pole. Science, 343(6167), pp.171-174.

Stillman, D. E., Michaels, T. I., Grimm, R. E., 2017. Characteristics of the numerous and widespread recurring slope lineae (RSL) in Valles Marineris, Mars. Icarus, 285, pp.195-210.

Thomas, E. C., Hodyss, R., Vu, T. H., Johnson, P. V., and Choukroun, M., 2017. Composition and evolution of frozen chloride brines under the surface conditions of Europa. ACS Earth and Space Chemistry, [e-journal] http://dx.doi.org/10.1021/acsearthspacechem.6b00003.

Van Hoolst, T., Rambaux, N., Karatekin, Ö, Dehant, V. and Rivoldini, A., 2008. The librations, shape, and icy shell of Europa. Icarus, 195(1), pp.386-399.

Using Viruses as Molecular Biology Tools: A Review of Viral Transneural Tracing

JACQUELINE WATT
Integrated Science Program, Class of 2017, McMaster University

SUMMARY

Viruses have been around for thousands of years, mastering invasion and evasion techniques of their host organisms. With the rapid progression of technology, it is these unique characteristics that can now be applied to studying cells and how they function. The most notable development is the use of neurotrophic viruses as transneural tracers. Transneural tracing traditionally involves the use of compounds that are able to pass through synapses in order to visualize connectivity between functional neurons. Neurotrophic viruses, specifically α-herpesvirus and rabies virus, have an innate ability to infect neurons and transfer between synapses. This property, along with their ability to self-amplify through replication in neurons, makes these viruses highly advantageous over traditional methods of transneural tracing. With the recent advancement of genetic engineering, there is great potential to combine genetic modification and viral transneural tracing, thus enabling more in-depth studies. This review aims to outline the unique characteristics of α-herpesvirus and rabies virus that makes them good candidates for transneural tracers and examine the potential that genetic modification can open for neuroscience research.

Received: 03/05/2017 **Accepted:** 03/30/2017 **Published:** 07/31/2017
URL: https://journals.mcmaster.ca/iScientist/article/view/1453

Keywords: transneural tracers, viruses, neural circuits, Herpesvirus, Rabies virus, viral neural transport, neurotrophic viruses, central nervous system, peripheral nervous system, motoneurons

INTRODUCTION

Manipulating and using viruses to benefit individuals and further the understanding of cells and how they function is not a new idea. Viruses are obligate, intracellular parasites that have evolved to survive within almost every species. Therefore, they provide a great opportunity to deliver treatment or study multiple species in vivo (Ugolini, 2010). This concept is especially useful when considering difficult systems to access, such as the nervous system. A revolutionary advancement in neuroscience research was the development of neurotrophic viruses as transneural tracers.

A transneural tracer is a tracer that is only able to pass through synaptically connected neurons, thus allowing for visualization of functional neurons (Callaway, 2008; Ugolini, 2010). For a tracer to be reliable there are certain requirements that must be met (Ugolini, 2010):

1. The tracer must only be able to transfer between neurons through synaptic connections.

2. The tracer should only travel in one direction through the axon.

3. The tracer must be able to label higher order neurons, meaning to pass through more than one synapse, in a neural circuit.

4. The number of synaptic connections the tracers passes through must be easily determined.

5. The tracer must be easily detectable and not degrade with time.

6. The tracer should not alter neural function.

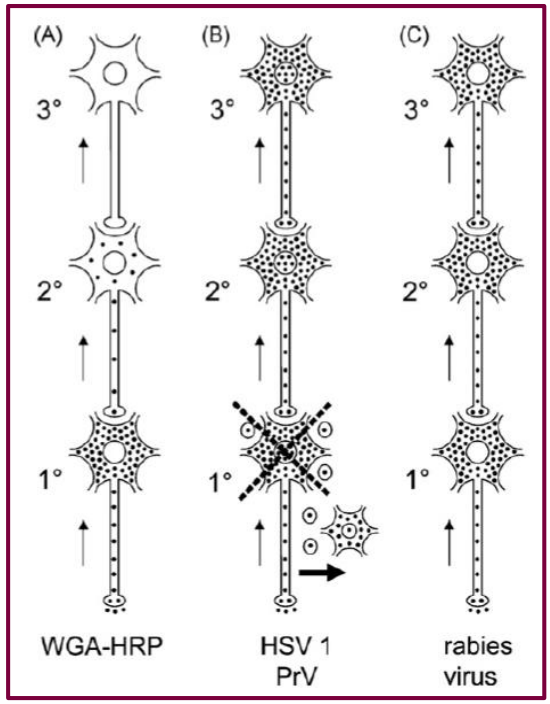


Figure 1: Diagram comparing the differences between transneural labelling with (A) conventional tracers, (B) α-herpesvirus, and (C) rabies virus. With conventional tracers only a low concentration of tracer is transferred to higher order neurons making them difficult to visualize. With viral transneural tracers a high concentration is maintained through each synaptic transfer allowing easy visualization of higher order neurons. With α-herpesvirus spurious spread is also possible (Ugolini, 2010).

Early non-viral transneural tracers, referred to as conventional tracers, included compounds that were stainable and could pass through synapses. Some examples of these compounds include horseradish peroxidase and dextran amines (Taber, et al., 2005). These compounds are weak transneural tracers as only a low concentration ever successfully crosses a synapse and consequently, the concentration is too low to label higher order neurons (Figure 1) (Ugolini, 2010). Therefore, neural circuits have to be inferred from multiple experiments showing individual synaptic connections, which is highly inefficient (Taber, et al., 2005).

Neurotrophic viruses are viruses that can enter and infect neurons. They have a natural ability to cross chemical synapses and are able to self-amplify (replicate) within each infected neuron (Callaway, 2008; Ugolini, 2010). These unique characteristics make them a much more sensitive and reliable transneural tracer, making them better than conventional tracers (Figure 1). Recently α-herpesvirus and rabies virus have been developed as transneural tracers, each with its own unique properties (Callaway, 2008; Ugolini, 2010). The purpose of this review is to discuss the properties of each of these viruses that make them useful for transneural tracing and highlight recent studies that were only successful due to these viral transneural tracers. The emerging technology of genetically modified viruses will also be discussed.

α-HERPESVIRUS

The first neurotrophic virus to be developed for transneural tracing was the α-herpesvirus, specifically Herpes Simplex virus type 1 (HSV1) and Pseudorabies virus (PrV) (Ugolini, 2010). HSV1 is able to infect and enter the central nervous system (CNS) from any injection site in all mammals, except for primates. To enter the CNS in primates, HSV1 must be injected directly into neurons of the CNS (Ugolini, 2010). PrV is a swine α-herpesvirus that can only be used as a transneural tracer for rodents and carnivores (Babic, et al., 1993 and Mettenleiter, 2003).

When considering using a virus as a transneural tracer, the virus-host interactions are crucial to the success of the study. These include what species are susceptible and permissible to the virus, how the age of the host effects the pathogenesis of the virus, and what dose and strain of the virus will be required to ensure the virus infects the desired cells. The α-herpesvirus was a good candidate for the first viral neural tracer as it is well characterised, thus these factors could be easily determined and controlled. Additionally, the number of synaptic connections can be determined using the kinetics of propagation of the virus. This would take into account the time of travel through the axon and the time to complete one replication cycle. The measurement interval for the visualization of one synaptic transfer can then be calculated for each specific strain (Ugolini, 2010). The virus is visualized using immunolabeling with florescent proteins such as green florescent protein (GFP) or red florescent

protein (RFP) (Enquist, et al., 2002). Studies have found that α-herpesvirus infection of neurons for transneural tracing results in neural degradation, thus must be limited to the onset of symptoms for ethical reasons (Ugolini, 2010).

VIRAL CHARACTERISTICS

α-herpesvirus is a large, double stranded DNA virus (Figure 2). It has a large susceptible host range; specifically, for HSV1 and PrV, humans and swine, respectively are the natural hosts, where latency is common (Tomishima, et al., 2001).

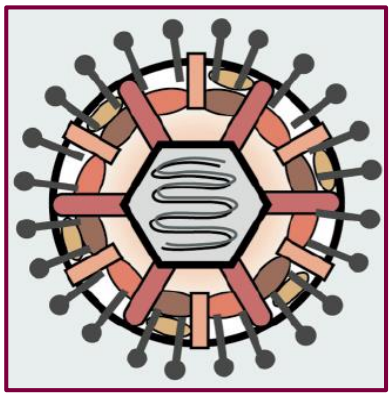


Figure 2: α-herpesvirus structure. All α-herpesvirus have a similar structure containing an icosahedral capsid, tegument, and membrane glycoproteins. The glycoprotein (gD) are illustrated as spikes on the envelope of the virion and are the part recognized by the host cell's receptors (Tomishima, et al., 2001).

α-herpesvirus enters the body through epithelial cells and then travels to the nucleus where the viral genome is released (Tomishima, et al., 2001). After initial infection, in natural hosts of the virus, α-herpesvirus can enter and remain dormant in dorsal root ganglia of the peripheral nervous system (PNS) (Enquist, et al., 2002). In this state the lytic cycle of the virus follows an alternative transcriptional program, mainly remaining transcriptionally inactive with limited expression of its genes (Ugolini, 2010). α-herpesvirus can persist in this state for the entire life of the host. However, the virus can be reactivated where it will then exit from the neuron and travel back to the original site of infection and begin actively transcribing again (Ugolini, 2010).

ENTRY INTO NEURONS

α-herpesvirus enters a neuron in the same manner as entering a cell, by plasma membrane fusion of its envelope (Tomishima, et al., 2001). The receptor for α-herpesvirus, which recognizes a glycoprotein (gD) on the outside of the virion, is speculated to be either from the tumor necrosis factor superfamily or similar to the nectin 1 and nectin 2 receptors of the immunoglobulin superfamily (Mettenleiter, 2003). From here, the capsid will travel in the axon, in retrograde motion, towards the cell body. The capsid moves along the microtubules in the axon, using the host cell's microtubule motors (Tomishima, et al., 2001). Due to the ability of the virus to reactivate, the capsid is also able to move in anterograde motion, away from the cell body, along the axon. The direction the capsid moves can be changed almost instantaneously, with the capsid having an equal probability for each direction of travel at any given time (Enquist, et al., 2002). Retrograde motion occurs approximately every one out of seven infections (Enquist, et al., 2002). This bi-directional capsid motility is an important characteristic of α-herpesvirus neural tracers as allows for varying tracing methods and techniques to be used.

Once at the cell body, the viral genes are released into the nucleus. Typically, the virus would establish latency at this point, however, when used as a transneural tracer, active replication occurs in the neuron. Assembly of the mature viral particles happens in the extra-nuclear compartments (Mettenleiter, 2003). At this point there are two pathways the virus can take (Figure 3). First, the whole virion can assemble in the extranuclear compartment. If this happens then the virus is more likely to travel in an anterograde direction along the axon (Mettenleiter, 2003). Second, the virus will not assemble into the complete viral particle and instead form sub-assemblies of the capsid, tegument, and envelope proteins (Mettenleiter, 2003). With the formation of sub-assemblies, retrograde motion along the axon is more probable, with the components accumulating at the synapse. Assembly of the whole viral particle would then occur near or at the axon terminus (Mettenleiter, 2003). The glycoprotein gE is believed to play a large role in initiating virion formation at the synapse, making it essential for transsynaptic transfer. The Us9 membrane protein, which is important for axonal localization of envelope proteins, determines

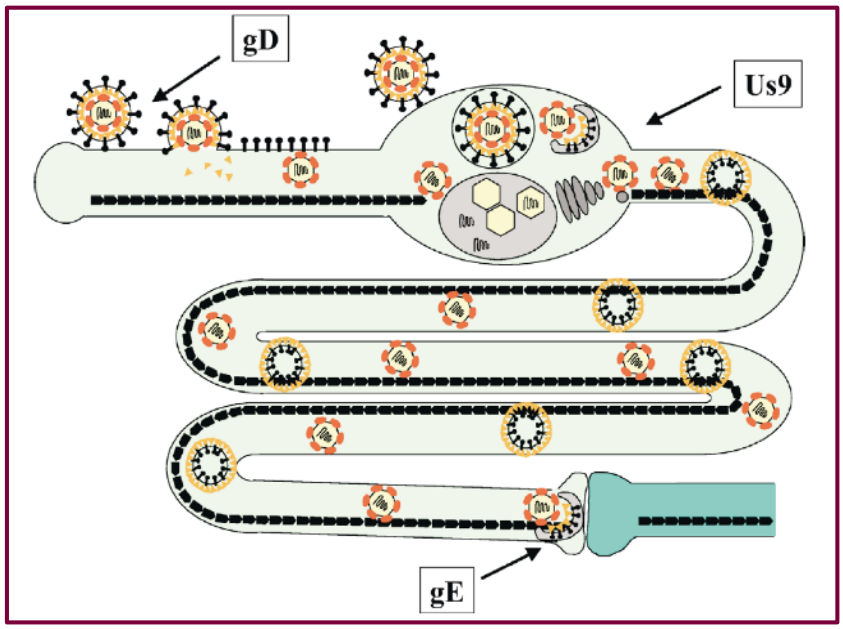


Figure 3: Summary of the possible events that can take place after an α-herpesvirus enters a neuron. gD is important for the virus to be recognized by the host cell upon initial infection. Us9 is responsible for determining which pathway the viral particle will follow after transcription. gE is essential for triggering virion assembly at the axon terminus and synaptic transfer (Mettenleiter, 2003).

which path the virus will take (Mettenleiter, 2003; Tomishima, et al., 2001). Along with synaptic transfer of α-herpesvirus from a neuron, direct cell-to-cell spread to other cells in close proximity such as astrocytes and microglia is possible (Mettenleiter, 2003). This spurious spread greatly reduces the reliability of α-herpesvirus as a transneural tracer.

USE AS A TRANSNEURAL TRACER

There are many unique characteristics of α-herpesvirus that make it a revolutionary transneural tracer. One of the biggest advantages being its capability for bi-directional motion along the axon. Depending on the purpose of the study, a specific direction of travel can be selected for based on the strain of α-herpesvirus (Babic, et al, 1996). This selection is possible due to certain mutations of the gE protein found in the different strains that greatly increases the probability of travel in a certain direction (Ugolini, 2010). For example, attenuated Bartha PrV and McIntyre-B strain of HSV1 prefer to travel in the retrograde direction, whereas, H129 strain of HSV1 prefers to travel in anterograde direction (Hoover and Strick, 1993; Kelly and Strick, 1993 and Zemanick, Strick and Dix, 1991). Additionally, due to α-herpesvirus' ability to enter the PNS it is the viral transneural tracer of choice for studying autonomic innervation, such as those related to heart rate and breathing (Ugolini, 2010).

Despite all of these advantages, there are some crucial limitations of using α-herpesvirus as a transneural tracer. In a natural infection, α-herpesvirus would establish latency in neurons and not undergo active transcription, results in rapid neural degradation of the infected neurons and glial cells. This is caused by a large inflammatory response in the neurons, triggered by the break down of host mRNA, since the virus has not adapted to control the immune response in neurons like it has in epithelial cells (Laurent, Madjar and Greco, 1998 and Smith, Malik and Clements, 2005). These genes are crucial for viral replication; therefore, they cannot be deleted or manipulated to potentially reduce the inflammatory response (Ugolini, 2010). When using α-herpesvirus as a transneural tracer, it is important to design the study in a way that helps mitigate the limitations.

STUDYING OREXIN-CONTAINING NEURONS USING DUAL TRACING WITH ISOGENIC STRAINS OF α-HERPESVIRUS

A prime example of the benefits of using α-herpesvirus as a transneural tracer is seen in a study by Geerling, et al. (2003). The purpose of their study was to determine if orexin-containing neurons of the lateral hypothalamic region of the brain are connected to multiple systems controlled by the sympathetic nervous system, referred to as outflow systems, in rats. Orexin is a neuropeptide that has been shown to have regulatory properties in wakefulness and arousal. The main effect this study focused on was the flight or fight response, specifically, looking at the adrenal gland innervation and the heart, through the stellate ganglion. The authors attempted to elucidate the function and organization of these neurons was a double-tracing method with Barth PrV (Geerling, et al., 2003). This method involves using two isogenic strains of the virus, each tagged with a different protein for visualization. Geerling, et al. (2003) tagged one strain with GFP (BPrV-GFP)

and the other strain with β-galactosidase (BPrV-βgal), a protein that can be stained for easy visualization. Bartha PrV, a retrograde tracer, was selected for this study as they wanted to trace the neural pathway starting from the outflow system to the brain. In this situation, retrograde refers to the direction of travel towards the brain and anterograde would refer to motion away from the brain.

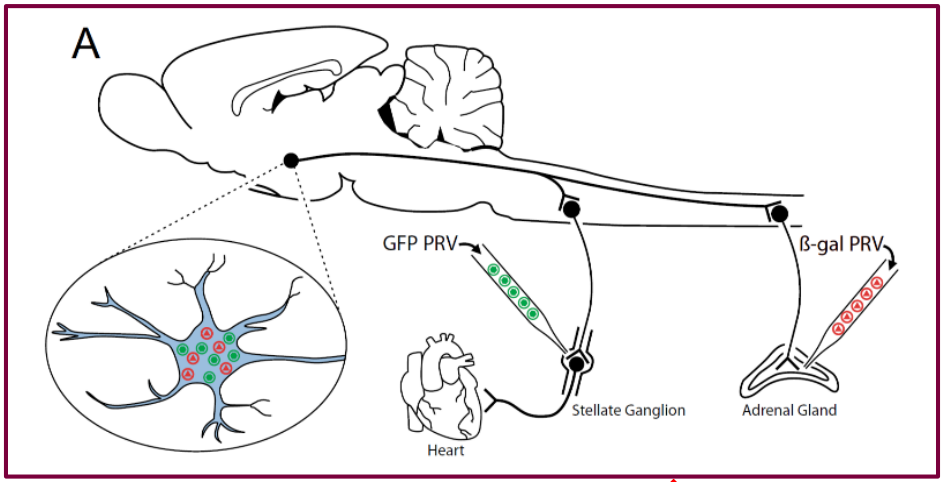


Figure 4: Design of the double tracing study. The GFP tagged strain of pseudorabies virus was injected into the stellate ganglion and the β-gal tagged strain was injected into the adrenal gland. The orexin neurons in the lateral hypothalamic region of the brain were then imaged to see if any contained the colours of both viral strains (Geerling, et al., 2003).

The BPrV-GFP strain was injected into one outflow system, the stellate ganglion. The BPrV-βgal strain was injected into the other outflow system, the adrenal glands (Figure 4). The virus was then allowed to replicate and synaptically transfer naturally over a period of four days, giving it more than enough time to travel up to the lateral hypothalamic region (Geerling, et al., 2003). This process all took place in vivo. If a neuron was infected with both strains, it had to have been synaptically connected to both sympathetic outflow systems, thus, have the possibility to control multiple targets. Geerling, et al. (2003) were able to show orexin-containing neurons that were tagged with both GFP and β-galactosidase (Figure 5).

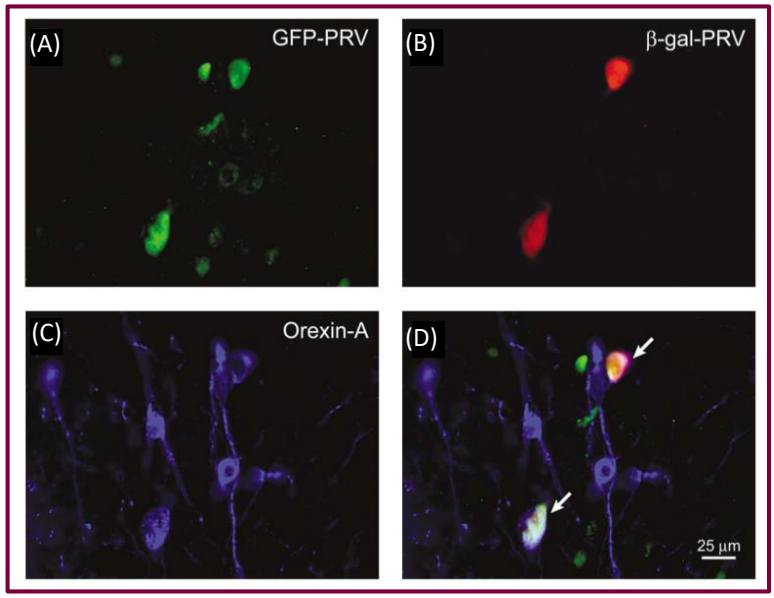


Figure 5: Images of the labelled orexin neurons in the lateral hypothalamic region after PrV was allowed time to replicate and transfer from the initial injection sites. (A) shows all neurons in the region that were infected with BPrV-GFP. (B) shows all neurons in the region that were infected with BPrV-βgal. (C) shows an image of the lateral hypothalamic region with a stain applied to show all orexin-containing neurons. (D) is an overlay of all the previous images. The white arrows indicate the neurons that exhibit colours from both viral strains (Geerling, et al., 2003).

Viral transneural tracing with BPrV was an essential component to the method of this study. The unique characteristics of the α-herpesvirus to enter the CNS from the PNS enabled the study of orexin-containing neurons in the lateral hypothalamic region of the brain. Additionally, the self-amplifying ability of the virus allowed it to cross multiple synapses while remaining at high enough concentrations to be visualized.

RABIES VIRUS

With the development of rabies as a retrograde transneural tracer, it quickly replaced the use of retrograde strains of α-herpesvirus (Ugolini, 2010). However, H129 remains the best option for studies requiring anterograde travel along the axon. Rabies is the only completely reliable and entirely specific transneural tracer currently available, based on the requirements previously discussed. Rabies causes no degradation of infected, meaning there are no restraints on the length of the study. Additionally, rabies can only be synaptically transferred and is not able to undergo direct cell-cell transfer, thus there is no risk of spurious spread (Ugolini, 2010). There are two strains of rabies available, the street strain and

the fixed strain. The fixed strain is approximately 100-10000 times less infectious than the street strain, thus in the context of viral neural tracing only the fixed strain is considered (Dietzschold, Schnell and Koprowski, 2005). Of the fixed strain, the CVS-11 subtype is most commonly used (Ugolini, 2010).

Rabies is able to infect and enter the CNS in all mammals, including primates, from any intramuscular injection site. The one major limitation of rabies is that it can only transfer to the CNS from motor neurons (Tang, et al., 2005). Similar to α-herpesvirus, rabies is a well characterized virus making it easy to determine and control all the necessary factors for successful transneural tracing, such as the required dose. Rabies replicates and travels through neurons in a time dependant manner, allowing the number of synaptic connections to be calculated using the kinetics of propagation of the virus, similarly to α-herpesvirus (Taber, et al., 2005). Rabies can also be easily visualized using immunolabeling with florescent proteins (Taber, et al., 2005).

VIRAL CHARACTERISTICS

Rabies is a small, non-segmented, negative stranded RNA virus with a characteristic bullet shape (Figure 6). Its genome contains up to ten genes, with only five being common between all strains of rabies virus (Schnell, et al., 2010). The natural host species of rabies is bats, however it has a large susceptible host range including insects, fish, mammals, reptiles, and crustaceans (Albertini, et al., 2012). Rabies enters the myocytes by receptor mediated endocytosis (Schnell, et al., 2010). It then quickly moves from myocytes into motor neurons at neuromuscular junctions and then on to the CNS. Once in the neuron the virus will travel to the cell body where genome replication and translation of viral proteins takes places (Schnell, et al., 2010). No

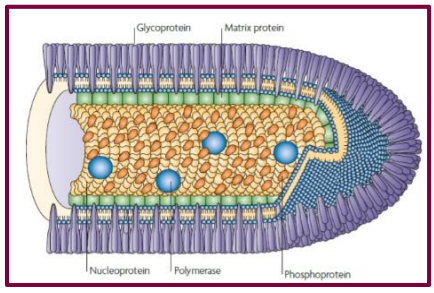


Figure 6: Rabies virus structure. All rabies viruses have a similar structure containing a neucleocapsid, envelope, and membrane glycoproteins (Schnell, et al., 2010).

host cell functions are manipulated or blocked with an initial rabies infection (Ugolini, 2010). These characteristics give rabies its uniquely long asymptomatic incubation period that ranges from 3 weeks to 3 months (Plotkin, 2000).

ENTRY INTO NEURONS

Rabies enters a neuron through receptor-mediated endocytosis, after recognition of its membrane glycoprotein G (Figure 7). While it is still uncertain which specific receptor binds to rabies, neural cell adhesion molecule (NCAM) is believed to play a role in its entry (Schnell, et al., 2010). Afterwards, one of two things are hypothesized to occur (Figure 8). First, the whole virion is transported along the axon in the vesicle from when it was endocytosed. Alternatively, membrane fusion of the virion and the endosome occurs soon after entry into the neuron and then the neucleocapsid is transported alone. Membrane fusion of the virus to the endosome is triggered by the low pH inside the vesicle. Current research more strongly supports the first theory (Albertini, et al., 2012).

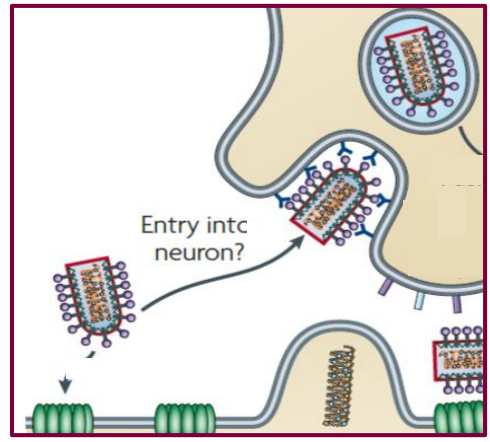


Figure 7: Rabies virus glycoprotein is recognized by the host cell and then enters by receptor mediated endocytosis (Schnell, et al., 2010).

Regardless of the method of transportation, the virus will travel in a retrograde motion, moving along the microtubules in the axon using the host's cytoplasmic dynein motor complex (Albertini, et al., 2012). Rabies can only travel in a retrograde motion, it is not capable of travelling in an anterograde direction along an axon. Once at the cell body, transcription and translation of the viral proteins begins. Very little is known about what

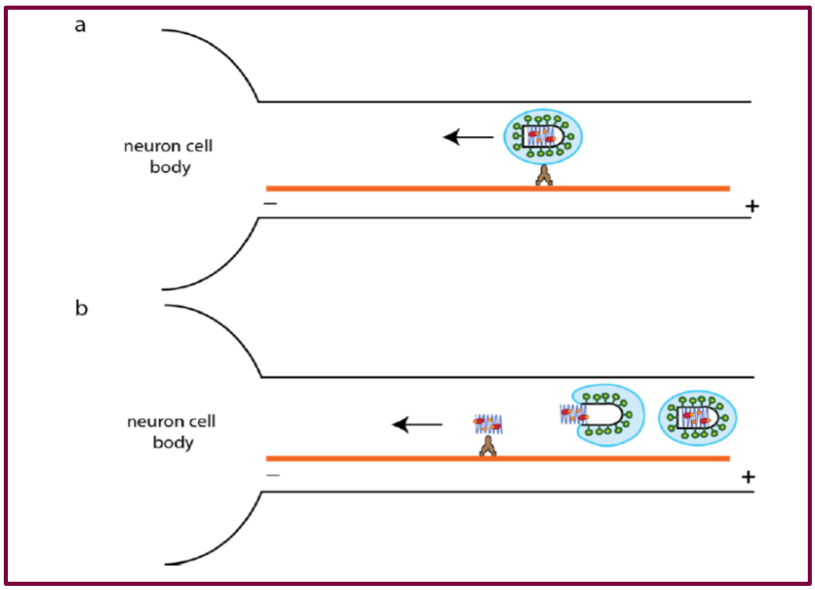


Figure 8: Two possible methods of axonal transport for rabies. (A) shows the whole virion moving along the axon inside a vesicle. (B) shows the early membrane fusion and the neucleocapsid moving along the axon on its own. Research more strongly favours the first method of transport (Albertini, et al., 2012).

triggers primary transcription (Albertini, et al., 2012). After replication, the virus is assembled and transported to the axon terminus. Unlike most neurotrophic viruses, rabies does not use apoptosis as a mechanism to spread in the body. Instead the mature virions will bud from the neuron and undergo synaptic transfer (Schnell, et al., 2010).

USE AS A TRANSNEURAL TRACER

The unique characteristics of rabies make it the most reliable transneural tracer currently available. This is due to the fact that it can only be transferred synaptically, with no spurious spread. There is also no neural degradation as a result of viral infection. Rabies does not induce apoptosis and has evolved to actively replicate in neurons and thus is able to control their immune response. The advantage of using the fixed strain is that it has been attenuated to the point where its incubation period and virulence are stabilized (Schnell, et al., 2010). The incubation period of the fixed strain is approximately one week. In that time period the virus can cross approximately seven synapses, which is more than enough transfers for a successful study (Ugolini, 2010).

There are a few limitations when using rabies as a transneural tracer. Since there is no neuronal degradation caused by infection, it can be difficult to determine the initial infection site. To overcome this limitation, dual tracer techniques can be used with a conventional tracer (Conti, Superti and Tsiang, 1986). The conventional tracer is only capable of spreading to first order neurons from the infection site, thus giving a more precise marker of the initial infection site. Cholera toxin B fragment, in a low concentration, is the only conventional tracer determined to effectively mark the infection site while not effecting the uptake of the rabies virus by the cells (Ugolini, 2010). The other limitation with rabies is the potential for some cytopathic changes (structural changes to the host cell due to a viral infection) in CNS neurons, however these changes are usually negligible (Jackson, 2002).

STUDYING A MODEL OF SCHIZOPHRENIA USING RABIES VIRUS

Technique wise, a notable example of using rabies as a transneural tracer is with a study by Brennand, et al. (2011). The aim of their study was to elucidate cellular and molecular defects caused by schizophrenia, specifically looking at neural connectivity. Schizophrenia (SCZD) is a debilitating neurological disease that affects approximately 1% of the world's population (Sullivan, Kendler and Neale, 2003). Despite its high prevalence, very little is known about the disease due to the difficulty of studying functional neurons in vivo. In order to overcome this problem, Brennand, et al. (2011) created schizophrenia human induced pluripotent stem cells (SCZD iPSC). iPSCs are adult cells that have been genetically reprogrammed to exhibit embryonic stem cell properties by overexpressing the genes and transcription factors that define the embryonic stem cell state (National Institutes of Health, 2015). Embryonic stem cells have the ability to differentiate into any cell of the body. Therefore, they can be directed to differentiate into neurons by expressing the correct mixture of signals and transcription factors normally

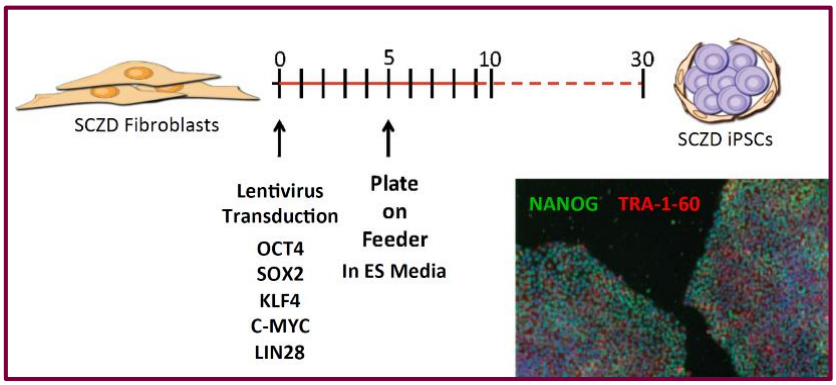


Figure 9: Design of the procedure for creating schizophrenia induced pluripotent stem cells (SCZD iPSC). The fibroblasts were collected from the schizophrenic patient then transduced with transcription factors that define the embryonic stem cell state. The cells were then cultured in a specific media and allowed to populate for 30 days (Brennand, et al., 2011).

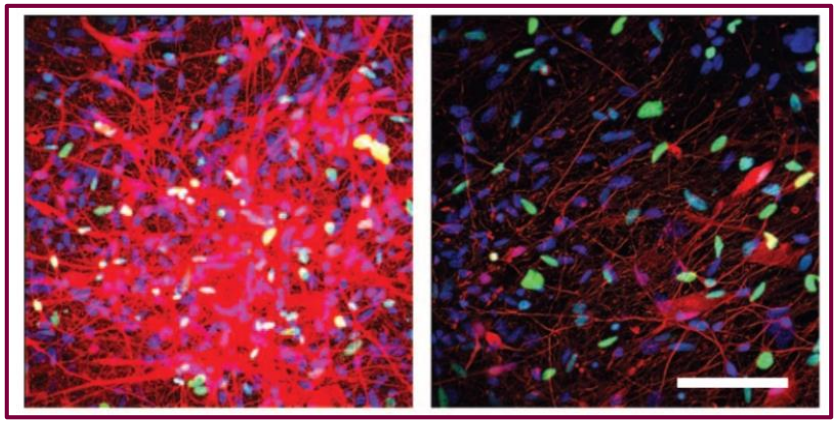


Figure 10: Images of the schizophrenia model after the rabies virus was allowed time to replicate and synaptically transfer. (A) is from an individual without schizophrenia. It shows a high concentration of red neurons indicating synaptic connection. (B) is from an individual with schizophrenia. It shows a low concentration of red neurons and more green neurons indicating very little synaptic connection (Brennand, et al., 2011).

expressed during neural development (National Institutes of Health, 2015).

Brennand, et al. (2011) took samples of and reprogrammed fibroblast cells, the most common cell found in connective tissues such as collagen, from schizophrenic patients. These reprogrammed cells were then cultured and later directed to differentiate into neurons (Figure 9). At the end of this process, Brennand, et al. (2011) had an accessible sample of functional neurons from a schizophrenic patient. Essentially, they created a model of the disease.

With their model they assayed for neural connectivity using rabies as a transneural tracer. First the primary neurons were transduced with GFP using a lentivirus. The rabies virus was then tagged with a RFP and injected into the model one week later (Brennand, et al., 2011). Primary infected cells were both GFP and RFP positive, while cells synaptically connected to the primary cells were GFP negative and RFP positive (Brennand, et al., 2011). Since rabies can only be transferred through synapses, this set up gave a clear indication of neural connectivity by comparing the number of GFP tagged neurons to RFP tagged neurons. After injecting rabies into the model, the virus was allowed to naturally replicate and propagate for 10 days before imaging the sample (Brennand, et al., 2011). This same procedure was then repeated for fibroblast cells from an individual without SCZD as a control.

After examining the images of both samples, Brennand, et al. (2011) were able to show that neurons in the SCZD model had significantly decreased connectivity than the control model (Figure 10). The complete reliability of rabies enabled them to quantify their results using the ratio of GFP tagged neurons to RFP tagged neurons. Additionally, they were able to test multiple antipsychotic drugs on the model and get a measurable result to determine if they had any significant effects on the neural connectivity. Brennand, et al. (2011) exposed both the control and SCZD models to five major antipsychotic drugs: loxapine, clozapine, olanzapine, risperidone, and thioridazine. Only loxapine significantly improved the neural connectivity (Figure 11).

This study would not have been successful before the development of rabies as a transneural tracer. The unique characteristic of rabies to be only synaptically transferred ensured the results were reliable and allowed for some cellular defects linked to SCZD to begin to be defined. Additionally, the use of a virus enabled the study of functional neurons, thus giving a more accurate understanding of the disease.

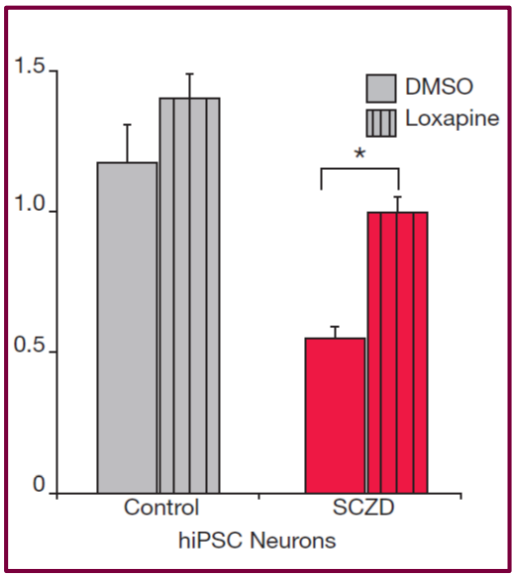


Figure 11: Graph showing the results of treating both control and schizophrenia models with loxapine. The vertical axis represents the ratio of RFP tagged neurons to GFP tagged neurons in each of the images. The solid coloured bar represents the negative control, treated with a solvent. The stripped bars represent the models treated with loxapine. The left, gray bars represent the control model. The right, red bars represent the schizophrenia model and show the significant improvement to neural connectivity after treatment. The error bars are standard error (Brennand, et al., 2011).

FUTURE OUTLOOK

GENETICALLY MODIFIED VIRAL TRANSNEURAL TRACERS

One of the most obvious advancements that could be made with viral transneural tracing is dual tracing with two different viruses in order to study broader systems and connectivity. However, infection of a cell with one virus interferes with the ability of another virus to infect the cell. The efficacy of the second virus to infect the cell exponentially decreases with time. The second virus must infect the previously infected cell within a short period of time, usually within an hour, of the first virus to have a chance for successful replication. This strict time dependant procedure also depends on the dose of each desired viral transneural tracer, making it difficult and tedious to perform. Therefore, researchers are focussed on discovering other more efficient areas of advancement.

The development of reverse genetics gave the ability to genetically manipulate viruses directly. Therefore, this enables viruses to be modified to fit a specific purpose or study and gives the potential for more accurate results while avoiding dual tracing with two different viruses (Callaway, 2008). α-herpesvirus has a large genome with many non-essential genes providing great possibility for genetic manipulation. Rabies has a small genome of only 10 genes at maximum, thus has less opportunity for manipulation. One of the main advantages of using genetically manipulated viruses is the ability to restrict the progression of the virus to a certain neural pathway starting from a specific cell.

Typically, after injecting a viral transneural tracer into a sample it will infect the cells non-specifically and spread through all of the connected synapses. A virus is not potent enough to have successful replication and tracing when injected into only a single target cell (Wickersham et al., 2007). This provides a great limitation to the specificity of studies especially considering similar neighbouring neurons can have drastically different connectivity and functions from each other. To overcome this problem, Wickersham, et al. (2007) designed a study using rabies virus that limited its infectivity

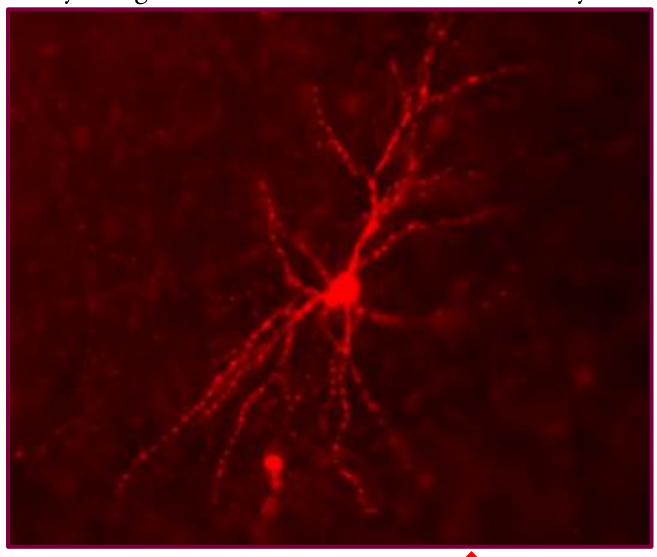


Figure 12: Image of a single stained neuron and its synaptic connections. A genetically modified rabies virus expressing envelope protein A and tagged with RFP was injected into a slice of brain from a rat and allowed to replicate. The target cell was transfected with sub-group A avian sarcoma receptor. The image was captured six days after initial infection showing infection of only the target neuron (Wickersham et al., 2007).

to a specific cell. The G protein, which is the part of the virus recognized by the cell for entry, was deleted. They then expressed the envelope protein A of the avian leukosis virus, which has the unique property of only being recognized by sub-group A avian sarcoma receptor, in the deletion mutant rabies virus (Wickersham, et al., 2007). The gene for the receptor was transfected into the target cell for their study. Therefore, the modified rabies virus would only be able to infect the cell containing the transfected avian receptor. With no other cells expressing the sub-group A avian sarcoma receptor, the rabies virus would be restricted to transsynaptic transfer after replication, therefore unambiguously tracing a specific neural pathway. Wickersham, et al. (2007) tested this procedure in a control experiment in vitro using rat brain slices to show proper infection selectivity. The images of their results showed that after injection of the genetically modified rabies virus into the sample rat brain slice, the virus successfully only infected the target cell (Figure 12).

Theoretically, another possibility for advancement is genetically modifying viruses in order to link connectivity to function in a single study. One of the main goals of neuroscience is to understand how the structure and function of neural circuits relate. Rabies virus recombinants could be designed to express neural activity indicators or light-sensitive ion channels (Ugolini, 2010). This would allow for the monitoring or manipulation of neural activity in vivo, which would be a major step forward in accomplishing the goal of fully understanding neural circuits. The use of rabies as a transneural tracer was essential to enable this emerging technique as it does not cause neural degradation in infected cells (Ugolini, 2010). The use of reverse genetics on viruses is in the early stages, however there is a great possibility that this research will lead to the next major step forward with understanding neural structure and function.

ACKNOWLEDGEMENTS

I would like to thank Dr. Sarah Gretton from the Natural Sciences programs at the University of Leicester for her supervision and guidance in completing this independent study.

REFERENCES

Albertini, A., Baquero, E., Ferlin, A. and Gaudin, Y. 2012. Molecular and cellular aspects of rhabdovirus entry. *Viruses*, 4, pp.117-139.

Babic, N., Klupp, B., Brack, A., Mettenleiter, T., Ugolini, G. and Flamand, A. 1996. Deletion of glycoprotein gE reduces the propogation of pseudorabies virus in the nervous system of mice after intranasal inoculation. *Virology*, 219, pp. 279-284.

Babic, N., Mettenleiter, T., Flammand, A. and Ugolini, G. 1993. Role of essential glycoproteins gII and gp50 in transneuronal transfer of pseudorabies virus from the hypoglossal nerves of mice. *Journal of Virology*, 67, pp. 4421-4426.

Brennand, K., Simone, A., Jou, J., Gelboin-Burkhart, C., Tran, N., Sangar, S., Li, Y., Mu, Y., Chen, G., Yu, D., McCarthy, S., Sebat, J. and Gage, F. 2011. Modelling schizophrenia using human induced pluripotent stem cells. *Nature Letters*, 473, pp.221-225.

Callaway, E. 2008. Transneuronal circuit tracing with neurotropic viruses. *Current Opinion in Neurobiology*, 18, pp.617-623.

Conti, C., Superti, F. and Tsiang, H. 1986. Membrane carbohydrate requirement for rabies virus binding to chicken embryo related cells. *Intervirology*, 26, pp. 164-168.

Dietzschold, B., Schnell, M. and Koprowski, H. 2005. Pathogenesis of rabies. *Current Topics in Microbiology and Immunology*, 292, pp. 45-56.

Enquist, L., Tomishima, M., Gross, S. and Smith, G. 2002. Directional spread of an α-herpesvirus in the nervous system. *Veterinary Microbiology*, 86, pp.5-16.

Geerling, J., Mettenliter, T. and Loewy, A. 2003. Orexin neurons project to diverse sympathetic outflow systems. *Neuroscience*, 122, pp.541-550.

Hoover, J. and Strick, P. 1993. Multiple output channels in the basal ganglia. *Science*, 259, pp. 819-821.

Jackson, A., 2002. Pathogenesis. In: Jackson, A. and Wunner, W. ed. 2002. *Rabies*. San Diego: Academic Press. pp. 245-282.

Kelly, R. and Strick, P. 2003. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of Neuroscience*, 23, pp. 8432-8444.

Laurent, A., Madjar, J. and Greco, A. 1998. Translational control of viral and host protein synthesis during the course of herpes simplex virus type 1 infection: evidence that initiation of translation is the limiting step. *Journal of General Virology*, 79, pp. 2765-2775.

Mettenleiter, T. 2003. Pathogenesis of neurotrophic herpesvirus: role of viral glycoproteins in neuroinvasion and transneuronal spread. *Virus Research*, 92, pp.197-206.

National Institutes of Health, 2015. *Stem Cell Basics*. [online] Available at: <http://stemcells.nih.gov/info/basics/pages/basics10.aspx> [Accessed 30 April 2016].

Plotkin, S. 2000. Rabies. *Clinical Infectious Diseases*, 30, pp. 4-12.

Schnell, M., McGettigan, J., Wirblich, C. and Papaneri, A. 2010. The cell biology of rabies virus: using stealth to reach the brain. *Nature Reviews Microbiology*, 8, pp.51-61.

Smith, R., Malik, P. and Clements, J. 2005. The herpes simplex virus ICP27 protein: a multifunctional post-translational regulator of gene expression. *Biochemical Society Transactions*, 33, pp. 499-501.

Sullivan, P., Kendler, K; and Neale, M. 2003. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Archives General Psychiatry*, 60, pp. 1187-1192.

Taber, K., Strick, P. and Hurley, R. 2005. Rabies and the cerebellum: new methods for tracing circuits in the brain. *The Journal of Neuropsychiatry & Clinical Neurosciences*, 17(2), pp.133-138.

Tang, Y., Rampin, O., Giuliano, F. and Ugolini, G. 1999. Spinal and brain circuits to motoneurons of the bulbospongious muscle: retrograde transneural tracing with rabies virus. *Journal of Comparative Neurology*, 414, pp. 167-192.

Tomishima, M., Smith, G. and Enquist, L. 2001. Sorting and transport of alpha herpesvirus in axons. *Traffic*, 2, pp.429-436.

Ugolini, G. 2010. Advances in viral transneuronal tracing. *Journal of Neuroscience Methods*, 194, pp.2-20.

Wickersham, I., Lyon, D., Barnard, R., Mori, T., Finke, S., Conzelmann, K., Young, J. and Callaway, E. 2007. Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. *Neuron*, 53, pp.639-647.

Zemanick, M., Strick, P. and Dix, R. 1991. Direction of transneural transport of herpes simplex virus 1 in the primate motor system is strain-dependant. *Proceeding of the National Academy of Science*, 88, pp. 8048-8051.

The Effects of Salinity and Activated Charcoal on the Herbivory of *Arabidopsis thaliana* by *Myzus persicae*

SOMMER CHOU, ALEXI DOAN, HELENA KONIAR, BRYCE NORMAN
Integrated Science Program, Class of 2019, McMaster University

SUMMARY

Road salt is commonly applied in the winter and inevitably percolates into surrounding areas where it is absorbed by plants. The detrimental effects of salinity on plants have been studied extensively, with recent literature investigating potential mitigative methods, including the application of biochar. This study explores the use of activated carbon, a modified version of charcoal with increased adsorptive ability, as a remediation technique for salt stress. Using model organisms *Arabidopsis thaliana* (thale cress) and *Myzus persicae* (green peach aphids), the aim was to determine the individual and combined effects of salinity and activated charcoal on plant performance and aphid populations using a factorial design. Overall findings presented a statistically significant effect (p=0.0201) on *M. persicae* herbivory between 25 mM salt and activated charcoal treatment groups. Although changes in plant biomass were not observed, there were a greater number of aphids occupying the plants without activated charcoal than on plants with activated charcoal for 25 mM salt treatments. Therefore, activated charcoal presents the opportunity for an accessible method of treatment for salt-stressed plants.

Received: 03/05/2017 **Accepted:** 03/30/2017 **Published:** 07/31/17
URL: https://journals.mcmaster.ca/iScientist/article/view/1454

Keywords: *Arabidopsis thaliana, Myzus persicae,* phloem-feeding insects, activated charcoal, road salt, plant-herbivore interactions, plant stress

INTRODUCTION

Road salt is universally applied to delay the formation of ice by lowering the freezing temperature of water. This application significantly reduces vehicular accidents in the winter (Turunen, 1997). However, its excessive use also entails severe, and possibly irreversible, environmental consequences for surrounding ecosystems (Moghaieb, Saneoka and Fujita, 2004; Parida and Das, 2005). Some countermeasures must be implemented to protect environmental integrity.

Without protective measures, road salt can have severe environmental consequences. Approximately 90% of road salt percolates as far as 6 metres from the road's edge as the snow melts, increasing the salinity of surrounding soil and groundwater (D. Williams, N. Williams and Cao, 2000; Norrström and Bergstedt, 2001). Road salt also interferes with various soil properties, such as cation exchange capacity, by leaching vital minerals involved in photosynthesis and cell structure maintenance out of the soil (Demarty, Morvan and Thellier, 1984; Berkowitz and Wu, 1993; Akselsson, et al., 2013). This leaching effect substantially decreases overall

plant performance, which in turn affects the entire ecosystem (Maron and Crone, 2006).

Plants are the basis of most terrestrial ecosystems as they provide energy and nutrients for herbivores (Bowdish and Stiling, 1998). When plant performance decreases, the resulting effects reverberate throughout the entire ecosystem and cause changes in both herbivory patterns and herbivore populations (Hairston, Smith and Slobodkin, 1960; Bowdish and Stiling, 1998). The application of various countermeasures to areas at risk of road salt exposure helps to avoid these detrimental effects. Studies have investigated the efficacy of protective methods, including silicon and proline application (Zhu, et al., 2004; Hoque, et al., 2007). A more recent study by Thomas, et al. (2013) also found that biochar can mitigate the detrimental effects of road salt. However, the impacts of activated charcoal, a substance that closely resembles biochar, on reducing salt toxicity are not entirely understood (Thomas, et al., 2013). Activated charcoal, otherwise referred to as activated carbon, is prepared through the pyrolysis of charcoal at temperatures between 950°C and 1000°C. This process stacks carbon atoms randomly, leaving interstices, or pores. This structural arrangement increases the overall surface area and thus increases the adsorptive ability of activated charcoal. This heightened adsorption is often exploited to remove unwanted tastes and smells as well as undesired organic and inorganic matter (Bansal and Goyal, 2005).

This study, conducted in September 2016, aims to determine how activated charcoal and salt water solutions influence aphid herbivory. To address whether activated charcoal reduces the impact of varying levels of soil salinity, we experimentally simulated a ditrophic ecosystem. *Arabidopsis thaliana* plants were inoculated with *Myzus persicae*, and treated with combinations of salt and activated charcoal. *A. thaliana* was selected as a model organism due to its relatively fast growth rate and *M. persicae* was selected due to its high fecundity (Baugh and Phillips, 1991; Somerville and Koornneef, 2002). In stressful conditions, *M. persicae* gives birth to winged alates, which can be interpreted as an indication of treatment-induced plant stress (Baugh and Phillips, 1991). Our experiment aims to answer three research questions:

1. Does varying salt concentration have an effect on the herbivory of *A. thaliana* by *M. persicae*?
2. Does activated charcoal have an effect on the herbivory of *A. thaliana* by *M. persicae*?
3. Does the interaction between activated charcoal and varying salt concentration affect the herbivory of *A. thaliana* by *M. persicae*?

MATERIALS & METHODS

To determine the effect of activated charcoal and salt concentration on aphid herbivory, a factorial design was generated involving 24 randomly selected wild type *A. thaliana* in the rosette stage of development. The McMaster Biology Greenhouse in Hamilton, Ontario, supplied all plants used in the study. *A. thaliana* plants were then partitioned into six groups of four, and each division was assigned an identifying letter corresponding to treatment received, as indicated in Table 1.

Table 1: Experimental Groups of *Arabidopsis thaliana*

	No Salt	Low Salt	High Salt
No Activated Charcoal	A	C	E
With Activated Charcoal	B	D	F

Indicators of plant performance such as number of leaves, rosette diameter, and stem length, were measured on day 1 prior to aphid inoculation and treatment administration. The number of leaves included all leaves visibly attached to the plant by a petiole, and excluded new buds. Rosette diameter was considered to be the longest distance from one leaf tip of the plant to the other, without extending any curled leaves. Stem length was measured as the length of the entire bolting stem extending from the top of the rosette, and was recorded as zero for non-bolting plants. Since all *A. thaliana* were initially selected to be in the rosette stage, each plant's day 1 stem length measurement was zero. These quantities were measured again on days 5, 8, and 12.

Activated charcoal was supplied by the McMaster Biology Greenhouse. Once all initial plant measurements and observations were recorded, activated charcoal was added to plants in groups B, D, and F (Table 1). A thin layer of activated charcoal, approximately 1-2 mm thick, was applied to the soil surrounding the plant on day 1 only.

Next, saline solutions were prepared by dissolving 0.3652 g and 0.7305 g of NaCl in 250 mL of distilled water to obtain the desired concentrations of 25 mM and 50 mM, respectively. These concentrations were chosen on the basis that *A. thaliana* can tolerate up to 100 mM of NaCl (Zhu, 2000). Plants that were not receiving saline treatments were watered with distilled water. Each plant was given a 20mL solution of their designated treatment on days 1, 5, 8, and 12.

M. persicae were raised on either *A. thaliana* or tobacco plants in the McMaster Biology Greenhouse. The aphids we used were in varying life stages, excluding winged alates. After watering, each individual plant was inoculated with three aphids. They were placed on the plant towards the middle of the rosette with a metal inoculating rod. On days 5, 8, and 12, the number of living aphids was recorded.

Plants were housed in two plastic trays throughout the duration of the experiment. To limit aphid transfer between plants to only winged-alates, paper dividers were positioned between plants, as seen in Figure 1.

Figure 1: Experimental design of *A. thaliana* trays. There were two trays with 12 plants each, separated by dividers to limit *M. persicae* movement to winged alates.

After watering and inoculation, plants were evenly distributed amongst the trays using a block-randomized method to minimize the effects of confounding variables. The two plant trays were divided in half for a total of four blocks (Figure 2). One plant of each experimental group (A-F) was included in each block. Additionally, the placement of plants within each block was randomized (Table 2). The trays were covered with a plastic lid containing two mesh-covered ventilation openings, and placed by a north-facing window for the duration of the study.

Block 1		Block 2	
D16	E20	D15	C12
A4	F21	E17	B5
B6	C11	F22	A1
Block 3		**Block 4**	
A3	D14	B8	C9
C10	B7	D13	A2
F24	E19	F23	E18

Figure 2: Summary of placement of each plant after block randomization. Numbers 1-24 refer to the specific plant, and letters A-F identify the treatments administered.

STATISTICAL ANALYSIS

The effect of salt concentration and activated charcoal along with their interaction on aphid herbivory was evaluated using a two-way analysis of variance (ANOVA) with interaction using statistical analysis software (Urbanek, Bibiko and Iacus, 2016). The data for number of aphids were log-transformed to improve homoscedasticity and normality, as aphids reproduce exponentially (Helms and Hunter, 2005). This also generated a better F critical value, indicating that the model was a better fit. The block effect due to the block randomization distribution of plants was accounted for in the ANOVA. Finally, Tukey's Honestly Significant Difference test was used to determine which interactions had significant variance in their means.

After the data was collected, one *A. thaliana* plant treated with low salt and no activated charcoal was removed due to an infestation of fungus gnat larvae. Any leaves touching the soil were wilted and irreparably damaged. This plant was an outlier within the data set and its poor performance was attributed to the fungus gnat larvae, leading us to exclude it in the analysis of variance.

RESULTS

Table 2: Analysis of variance results for effects of salt concentration, charcoal presence, and their interaction on the log-transformed population size of *M. persicae* on plants

Sources of Variation	Salt	Activated Charcoal	Salt & Activated Charcoal	Error
Degrees of Freedom	2	1	2	17
Mean Square	0.193	0.065	2.068	0.416
F Critical Value	0.4636	0.1556	4.9616	
p-value	0.6368	0.6981	0.0201	

As seen in Table 2, there was no significant effect of salt concentration on aphid population (p=0.6368) or activated charcoal on aphid population (p=0.6981). However, a significant effect was seen in the interaction between salt concentration and activated charcoal on aphid population (p=0.0201). This supports the hypothesis that the interaction between salt concentration and activated charcoal affects aphid herbivory. Thus, the effect of salt concentration on aphid herbivory is dependent on the presence or absence of activated charcoal and vice versa.

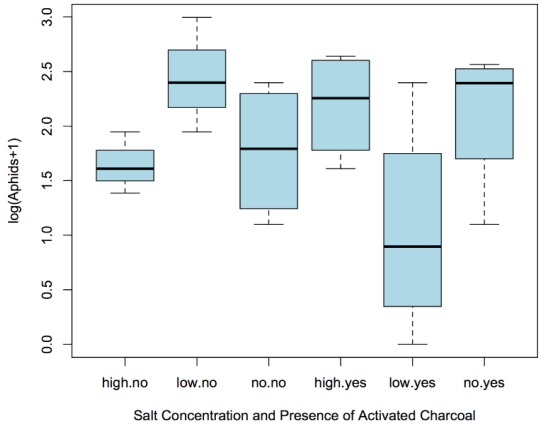


Figure 3: Mean value of (log_{10}+1)-transformed number of aphids on treated plants. There is a significant difference between means for low salt with activated charcoal and low salt without activated charcoal ($p<0.05$).

Figure 3 illustrates the mean value of log-transformed number of aphids on day 12 with their standard deviation for each treatment group. The difference in mean aphid populations is greatest for plants treated with a low salt concentration.

Figure 4 illustrates the mean value of log-transformed number of aphids on plants treated with no, low concentration, and high concentration salt water in either the presence or absence of activated charcoal. Plants treated with no salt or high salt had approximately the same number of aphids regardless of whether or not charcoal was present. Only groups of *A. thaliana* treated with low salt showed significant differences in number of aphids depending on the presence or absence of activated carbon.

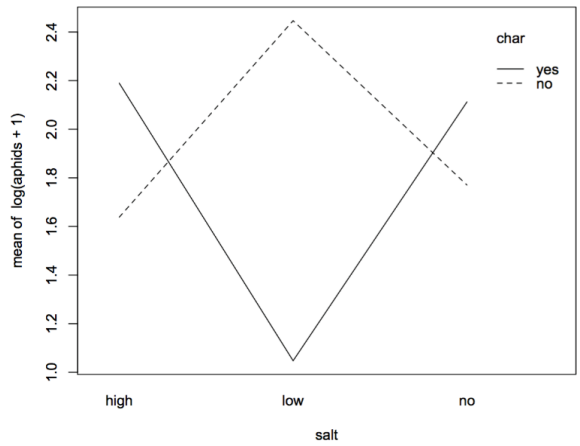


Figure 4: Interaction plot showing mean value of (log_{10}+1)-transformed number of aphids for each of the treatment groups. There is a significant difference in means of *M. persicae* on *A. thaliana* treated with low salt with and without activated carbon ($p<0.05$).

Salt concentration and activated charcoal did not have an effect on any measured plant biomass variables, such as number of leaves, stem length, or rosette diameter. The experiment was organized with a block random design to reduce the probability of confounding variables biasing the results. An analysis of variance showed the blocks had no significant effect on aphid herbivory.

DISCUSSION

Road salt is necessary for safe winter driving, but can abiotically stress plants and negatively impact plant performance. Biochar has been researched as a countermeasure for salt-induced plant stress (Thomas, et al., 2013). However, the effects of activated charcoal on plant-herbivore relationships are not yet well understood. A further understanding of activated charcoal's effects provides the possibility for a protective

countermeasure against salt damage on sensitive plants. This study examined how the relationship between *M. persicae* and *A. thaliana* was affected by the introduction of varying concentrations of salt water and activated charcoal. A two-way ANOVA with interaction showed that the effect of plants treated with 25 mM salt concentration on aphid herbivory significantly depended on the presence or absence of activated charcoal. However, aphid herbivory for plants treated with no salt or 50 mM salt concentration did not depend on the presence or absence of activated charcoal. Additionally, salt concentration and activated charcoal did not independently affect aphid herbivory.

There were more aphids on plants treated with low salt in the absence of charcoal than on plants treated with low salt in the presence of charcoal. These results indicate that aphid herbivory was higher in the absence of activated charcoal (Figure 4), which supports the plant stress hypothesis. This hypothesis states that herbivores preferentially feed on plants that are under biotic and abiotic stressors over unstressed plants (Joern and Mole, 2005). Piercing-sucking insects, including *M. persicae*, perform better and have higher reproductive potential on stressed plants compared to non-stressed plants (Koricheva, Larsson and Haukioja, 1998). This pattern could occur for several reasons: the alteration of biochemical source-sink relationship allocations, a change in foliar chemistry, and decreased defense levels (Joern and Mole, 2005). The changes in biochemical source-sink relationship allocations affect internal resource partitioning within stressed plants (Joern and Mole, 2005). A study performed by Lincoln (1993) found that, on leaves with elevated carbon content, there was increased consumption by herbivorous insects. This was due in part to the newly enhanced leaf digestibility and reduced efficiency of nitrogen use (Lincoln, 1993). The introduction of stressors to plants also induces a change in foliar chemistry by affecting the levels of nitrogen-containing compounds available within the leaves. These compounds include amino acids, amides, and imino acids (Mansour, 2000). Herbivores preferentially feed on stressed plants with higher nitrogen concentrations, and tend to survive longer as a result (Mattson, 1980).

There was no significant effect on the amount of aphid herbivory for plants that were treated with distilled water, independent of whether or not charcoal was present. This indicates that activated charcoal alone does not affect herbivory. Regardless of activated charcoal's ability to mitigate the negative effects of toxic chemicals, we found no evidence that it improved plant performance. Similarly, there was no relationship between the plants belonging to the two groups that were treated with a 50 mM salt solution. It may be that the charcoal's adsorptive ability was not high enough to mitigate the effect produced by a greater salt concentration. Independent of charcoal's presence or absence, when the salt concentration is 50 mM aphids are equally successful at reproducing on the salt-stressed plants. This suggests that the effect of activated charcoal on aphid herbivory depends on the level of salt concentration.

Aphids on *A. thaliana* stressed with 25 mM salt water without charcoal performed better than aphids on plants stressed with high concentration salt water without charcoal (Figure 4). This agrees with the pulsed stress hypothesis, proposed by Huberty and Denno (2004). This hypothesis states that herbivores, especially phloem feeders, perform better on plants that cycle through periods of stress and recovery (Huberty and Denno, 2004). Elongated periods of stress negatively affect the amount of nutrients that phloem-feeding insects can extract from a plant, while short cyclic periods maintain nutrient availability while decreasing plant performance (Huberty and Denno, 2004). Since the experiment had a relatively short time span (12 days), *A. thaliana* treated with low salt without charcoal experienced a stress period that allowed for increased performance of *M. persicae*. Plants given high salt without charcoal had a lower number of aphids compared to those exposed to low salt without charcoal due to the increased stress of *A. thaliana.* This reduced the nutrients available to *M. persicae* therefore causing a decrease in population. If the experiment were to be conducted over a longer period of time, plants treated using low salt without charcoal would be under an elongated period of stress, likely producing conditions less favourable for aphid herbivory.

The study results indicate that the presence or absence of activated charcoal alone did not have an effect on aphid herbivory. This is due to the fact that activated charcoal works to adsorb harmful toxins rather than to provide the plants with nutrients (Bansal and Goyal, 2005). It would appear that activated charcoal does not help nor hinder plant performance or independently affect aphid herbivory.

There are, however, innate limitations to this study's design. Within each of the 6 treatment groups, there were only 4 individual *A. thaliana*. This sample size is extremely small, due to resource restrictions. In small sample sizes, the presence of one outlier has more influence on prevalent trends within the data than in larger sample sizes. Since there were only 4 plants within each treatment group, having one outlier meant that a quarter of the data collected from that treatment group must be disregarded. Also, this experiment was performed over the course of 12 days, which is a very short time frame relative to *A. thaliana*'s life span of 2.5 months (Tocquin, et al., 2003). With a longer experimental time frame, the results would better represent changes in rosette diameter, number of leaves, and stem length. There was not enough time to allow for plant biomass to significantly accumulate or diminish over the course of the experiment, which likely contributed to the insignificance of plant performance results.

The interaction between activated charcoal and salt water should be examined in greater detail. Activated charcoal could be considered as a method for protecting sensitive environments to alleviate plant stress and decrease insect herbivory. This study did not analyze the effects of charcoal on organisms occupying other trophic levels. Future research could investigate the possible detrimental effects of activated carbon on different species within the food chain, a factor that could not be tested due to the ditrophic nature of our experiment.

The results of this experiment have several future implications. This gives insight into the efficacy of applying activated charcoal in tandem with road salt to minimize anthropogenic impacts and maintain road safety. Since our study only focused on the interaction between *A. thaliana* and *M. persicae*, the impact on a more complicated multitrophic ecosystem can be extrapolated from our simple design.

CONCLUSION

The interaction of 25 mM salt solutions with activated charcoal has a significant effect on the herbivory of *M. persicae* with an increased number of aphids on plants growing in the absence of activated charcoal. These results support the plant stress hypothesis and provide insight for a new method of protection for salt-stressed plants. With further research, activated charcoal could be widely applied to decrease insect herbivory and mitigate salinity effects on plants in heavily salted areas.

ACKNOWLEDGEMENTS

We would like to thank Dr. Susan Dudley for her guidance throughout the study and suggestions on the paper. We also thank Sebastian Irazuzta, Geneviève van Wersch, and Russ Ellis for their supervision during data collection days. Materials for the study were provided by the McMaster Biology Greenhouse in Hamilton, Ontario.

REFERENCES

Akselsson, C., Hultberg, H., Karlsson, P.E., Pihl Karlsson, G. and Hellsten, S., 2013. Acidification trends in south Swedish forest soils 1986–2008 — Slow recovery and high sensitivity to sea-salt episodes. *Science of The Total Environment*, 444, pp.271–287.

Bansal, R.C. and Goyal, M., 2005. *Activated Carbon Adsorption*. Boca Raton, FL: Taylor & Francis.

Baugh, B.A. and Phillips, S.A., 1991. Influence of Population Density and Plant Water Potential on Russian Wheat Aphid (Homoptera: Aphididae) Alate Production. *Environmental Entomology*, 20(5), p.1344 LP-1348.

Berkowitz, G.A. and Wu, W., 1993. Magnesium, potassium flux and photosynthesis. *Magnesium research*, 6(3), pp.257–265.

Bowdish, T.I. and Stiling, P., 1998. The influence of salt and nitrogen on herbivore abundance: direct and indirect effects. *Oecologia*, 113(3), pp.400–405.

Demarty, M., Morvan, C. and Thellier, M., 1984. Calcium and the cell wall. *Plant, Cell & Environment*, 7(6), pp.441–448.

Hairston, N.G., Smith, F.E. and Slobodkin, L.B., 1960. Community Structure, Population Control, and Competition. *The American Naturalist*, 94(879), pp.421–425.

Helms, S.E. and Hunter, M.D., 2005. Variation in plant quality and the population dynamics of herbivores: there is nothing average about aphids. *Oecologia*, 145(2), pp.196–203.

Hoque, M.A., Okuma, E., Banu, M.N.A., Nakamura, Y., Shimoishi, Y. and Murata, Y., 2007. Exogenous proline mitigates the detrimental effects of salt stress more than exogenous betaine by increasing antioxidant enzyme activities. *Journal of Plant Physiology*, 164(5), pp.553–561.

Huberty, A.F. and Denno, R.F., 2004. Plant Water Stress and its Consequences for Herbivorous Insects: A New Synthesis. *Ecology*, 85(5), pp.1383–1398.

Joern, A. and Mole, S., 2005. The Plant Stress Hypothesis and Variable Responses by Blue Grama Grass (Bouteloua gracilis) to Water, Mineral Nitrogen, and Insect Herbivory. *Journal of Chemical Ecology*, 31(9), pp.2069–2090.

Koricheva, J., Larsson, S. and Haukioja, E., 1998. Insect Performance on Experimentally Stressed Woody Plants: A Meta-Analysis. *Annual Review of Entomology*, 43(1), pp.195–216.

Kuemmel, D. and Hanbali, R., 1992. *Accident Analysis of Ice Control Operations Final Report*. [online] Milwaukee. Available at: <http://www.trc.marquette.edu/publications/IceControl/ice-control-1992.pdf> [Accessed 23 Oct. 2016].

Lincoln, D.E., 1993. The influence of plant carbon dioxide and nutrient supply on susceptibility to insect herbivores. *Vegetatio*, 104(1), pp.273–280.

Mansour, M.M.F., 2000. Nitrogen Containing Compounds and Adaptation of Plants to Salinity Stress. *Biologia Plantarum*, 43(4), pp.491–500.

Maron, J.L. and Crone, E., 2006. Herbivory: effects on plant abundance, distribution and population growth. *Proceedings of the Royal Society B: Biological Sciences*, 273(1601), p.2575 LP-2584.

Mattson, W., 1980. Herbivory in Relation to Plant Nitrogen Content. *Annual Review of Ecology, Evolution, and Systematics*, 11(1), pp.119-161.

Moghaieb, R.E.A., Saneoka, H. and Fujita, K., 2004. Effect of salinity on osmotic adjustment, glycinebetaine accumulation and the betaine aldehyde dehydrogenase gene expression in two halophytic plants, Salicornia europaea and Suaeda maritima. *Plant Science*, 166(5), pp.1345–1349.

Norrström, A.-C. and Bergstedt, E., 2001. The Impact of Road De-Icing Salts (NaCl) on Colloid Dispersion and Base Cation Pools in Roadside Soils. *Water, Air, and Soil Pollution*, 127(1), pp.281–299.

Parida, A.K. and Das, A.B., 2005. Salt tolerance and salinity effects on plants: a review. *Ecotoxicology and Environmental Safety*, 60(3), pp.324–349.

Somerville, C. and Koornneef, M., 2002. A fortunate choice: the history of Arabidopsis as a model plant. *Nature Reviews Genetics*, 3(11), pp.883–889.

Thomas, S.C., Frye, S., Gale, N., Garmon, M., Launchbury, R., Machado, N., Melamed, S., Murray, J., Petroff, A. and Winsborough, C., 2013. Biochar mitigates negative effects of salt additions on two herbaceous plant species. *Journal of Environmental Management*, 129, pp.62–68.

Tocquin, P., Corbesier, L., Havelange, A., Pieltain, A., Kurtem, E., Bernier, G. and Périlleux, C., 2003. A novel high efficiency, low maintenance, hydroponic system for synchronous growth and flowering of Arabidopsis thaliana. *BMC Plant Biology*, 3(1), p.2.

Turunen, M., 1997. Measuring salt and freezing temperature on roads. *Meteorological Applications*, 4(1), pp.11–15.

Urbanek, S., Bibiko, H.J. and Iacus, S.M., 2016. R. (3.2.3). [computer program]. R Foundation for Statistical Computing.

Williams, D.D., Williams, N.E. and Cao, Y., 2000. Road salt contamination of groundwater in a major metropolitan area and development of a biological index to monitor its impact. *Water Research*, 34(1), pp.127–138.

Zhu, J.-K., 2000. Genetic Analysis of Plant Salt Tolerance Using Arabidopsis. *Plant Physiology*, 124(3), pp.941–948.

Zhu, Z., Wei, G., Li, J., Qian, Q. and Yu, J., 2004. Silicon alleviates salt stress and increases antioxidant enzymes activity in leaves of salt-stressed cucumber (Cucumis sativus L.). *Plant Science*, 167(3), pp.527–533.

Investigating the Relationship between Variance of Transition Temperatures in an Undergraduate Laboratory Setting

ANGELICO OBILLE
Integrated Science Program, Class of 2017, McMaster University

SUMMARY

Colligative properties are properties of impurities. Most substances are mixtures since it is difficult to maintain a completely pure substance. The effects of a solute on the properties of a substance include freezing point depression (FPD) and boiling point elevation (BPE). These properties are due to the impurity lowering the mole fraction of the solvent, consequently decreasing its chemical potential. This stabilizes the liquid state and causes the solvent to exist as a liquid for a wider range of temperatures at constant pressures. The purpose of this study is to investigate a general relationship between FPD and BPE. Specifically, this study aims to provide a method in determining a proportionality between the transition constants of a general solvent. The transition constants represent how much the variance of transition temperatures change against varying concentrations of solute. The transition constants (K_{fp} and K_{bp}) are determined for water with varying concentrations of dissolved acetic acid, sucrose, and sodium chloride. The freezing point and boiling point of the solutions is determined with analysis of cooling and heating curves, and the variance of transition temperature is determined upon comparison with pure controls. BPE has been found to be positively correlated with FPD, changing by a factor of 0.35 (± 0.09 s.d.). The experiments designed in this study provide a simple reproducible way of demonstrating and supporting the fundamental principles underlying colligative properties.

Received: 03/05/2017 **Accepted:** 04/01/2017 **Published:** 07/31/2017
URL: https://journals.mcmaster.ca/iScientist/article/view/1461

Keywords: colligative properties, freezing point depression, boiling point elevation, transition constants, lab practicum, pedagogy

INTRODUCTION

Colligative properties are exhibited as a change in physical properties of a solvent with a dissolved solute. Examples of colligative properties are freezing point depression (FPD), boiling point elevation (BPE), osmotic pressure depression (OPD), and vapour pressure depression (VPD). These properties depend simply on the ratio of the number of solute particles to the number of solvent molecules in a solution. In other words, the chemical reactions of the different substances within the solution are not important factors when investigating the colligative properties of solvents. As such, the significant value of measure that determines the effects of a solute on a solvent is the van't Hoff factor of substances. This value allows for colligative properties to be measured upon apparent concentrations of solute particles within solution rather than the molecular concentration. For example, dissolving 1 mole of sodium chloride (NaCl) in water yields an apparent concentration of 2 moles within solution since each molecule of this substance dissociates completely in water to produce 2 ions, Na^+ and Cl^-. Furthermore, a substance that does not dissociate into parts upon solvation has a van't Hoff factor of 1. These values do not necessarily have to be integers depending on the nature of the solute. For example, acetic acid (CH_3COOH) solvated by water has a van't

Hoff factor of neither 1 nor 2. Acetic acid is a weak acid that may dissociate into its conjugate base and a hydronium ion, which would suggest that it should have a van't Hoff factor of 2. However, since it is a weak acid, it does not dissociate completely, i.e., some molecules in the solution are not dissociated and would thus have a van't Hoff factor of 1. Therefore, acetic acid actually has a van't Hoff factor of 1.05. It is important to define and consider van't Hoff factors of solutes when observing colligative properties in order to properly draw conclusions upon investigation.

The liquid state is the mediating state of matter between the most rigid structures and the most free-flowing fluids. It is an important state of matter because of its properties of maintaining a dynamic system while remaining contained within a space. Chemists use the liquid phase because reactions can happen much more quickly and in a more controlled fashion within solution. This study looks at the stabilization of the liquid state as a result of solutes dissolved in the mixture. The depression of freezing point temperature and the elevation of boiling point temperature are the observed physical changes caused by the presence of solvated particles. The application of FPD is prevalent especially during winter as road salt is dissolved in the water to prevent it from freezing. Changing the concentration of dissolved solutes is positively correlated to changing the amount by which the transition temperature of the solvent depresses or elevates. In other words, a certain amount of solute dissolved in a solution causes the freezing point to decrease a certain amount and if a greater amount of that same substance is dissolved, then the freezing point would decrease a greater amount. This also applies to the elevation of boiling point. Finding a relationship between these two properties would allow for the techniques developed in this experiment to be used to demonstrate and support the fundamental principles underlying colligative properties.

This study aims to answer the question of whether or not the freezing point depresses at the same rate at which the boiling point elevates about the same changes in dissolved solute concentration. In order to answer this question, a reproducible method for determining each the freezing point and boiling point must be used. This study uses the analysis of cooling and heating curves that map out the temperature of the solution over time. Temperature is simply a measure of the average kinetic energy of the particles in the solution. Observing an increase in temperature upon heating a substance is due to the energy transfer converting the thermal energy from the heat source into kinetic energy in the substance. However, thermal energy introduced to a system can also allow for intermolecular bonds, such as hydrogen bonds between water molecules, to break. The state of matter of a system is determined by the amount of intermolecular bonds exhibited between molecules of the substance. Less intermolecular bonds are exhibited in a gas than in a liquid, and more intermolecular bonds are exhibited in a solid than in a liquid. In a transition of state, the thermal energy transfer is used for the breaking of intermolecular bonds in melting and vaporization, and for the making of bonds in condensation and freezing. Thus, there is no thermal energy transfer contributing to the kinetic energy of the system. This would be exhibited in a cooling or heating curve as a plateau in temperature as the substance is changing state. This plateau can be easily interpolated to determine the temperature at which the state transition occurs and can thus determine the freezing and boiling points of various solutions. FPD and BPE are then determined by calculating the difference between the transition temperature of the solution and the transition temperature of the pure solvent control. The rate of FPD and BPE about changing concentration of solute particles is quantified as the transition constants: K_{fp} for freezing point and K_{bp} for boiling point. These constants can be determined through linear regression on the variance of transition temperature interaction with solute concentration.

If there is a relationship between FPD and BPE, then there would be a proportionality between the transition constants. If FPD and BPE change at the same rate, the regressions should be collinear and the ratio of transition constants (K_{bp}/K_{fp}) would be equal to 1. If FPD and BPE do not change at the same rate, but are proportional to each other, this would be observed by comparing the K_{bp}/K_{fp} ratio for multiple solutes and having the proportionalities be equal.

MATERIALS AND METHODS

This study investigated the change in the freezing point and boiling point of water in the presence of three solutes: sodium chloride (i=2), acetic acid

(i=1.05), and sucrose (i=1). The freezing point and boiling point of distilled water were determined and recorded as controls. Solutions of 0.45 mol/kg, 0.90 mol/kg, and 1.2 mol/kg of each solute were prepared in distilled water and tested with three replicate aliquots. 30 mL of each solution were prepared and three aliquots of 4 mL were taken in Pyrex® test tubes to be tested. The freezing point of each solution was also determined and the same frozen solution was tested afterwards to determine the boiling point.

Each solution was taken to freezing by submergence in an ice bath prepared by dissolving sodium chloride salt in a Styrofoam calorimeter cup containing a mixture of ice and water. Each solution was then taken to boiling by heating upon a hot plate in a heating block to evenly distribute the thermal energy among the solution.

Temperature of each solution through the transition processes was recorded with a PASCO® PASPORT Absolute Pressure/Temperature Sensor (PS-2146) and cooling/heating curves were constructed by a PASCO® SPARK Science Learning System (PS-2008A) interface.

A total of 30 cooling curves and 30 heating curves were constructed to determine freezing point and boiling point for each aliquot: 3 control aliquots, and 3 aliquots for 3 concentrations (0.45 mol/kg, 0.90 mol/kg, and 1.2 mol/kg). FPD is determined by subtracting the new freezing point from the freezing point of the distilled water control. Similarly, BPE is determined by subtracting the new boiling point by the boiling point of the distilled water control.

It is important to take the freezing point and boiling point of the solvent in which the query solutes are dissolved in because theoretical values can be significantly affected by even small amounts of impurities. As such, although theoretically water has a freezing point of 0°C and a boiling point of 100°C, the freezing point and boiling point determined by the protocol are different since there may be ions suspended in the distilled water. Since colligative properties are relative properties, it does not matter whether or not the solvent is pure.

RESULTS

It was determined that the freezing point of the distilled water control was -0.18 ± 0.01°C and the boiling point of the distilled water control was 101 ± 1°C. All freezing point and boiling point comparisons were taken from these values. The variance of transition temperature (ΔT_{tr}) values were plotted against concentration to compare the rate of BPE and the rate of FPD for sodium chloride, acetic acid, and sucrose dissolved in distilled water, as shown in Figure 1. Thus, it has been shown with this method that there is a proportionality between FPD and BPE and it is found that boiling point elevates at a rate slower than the depression of freezing point.

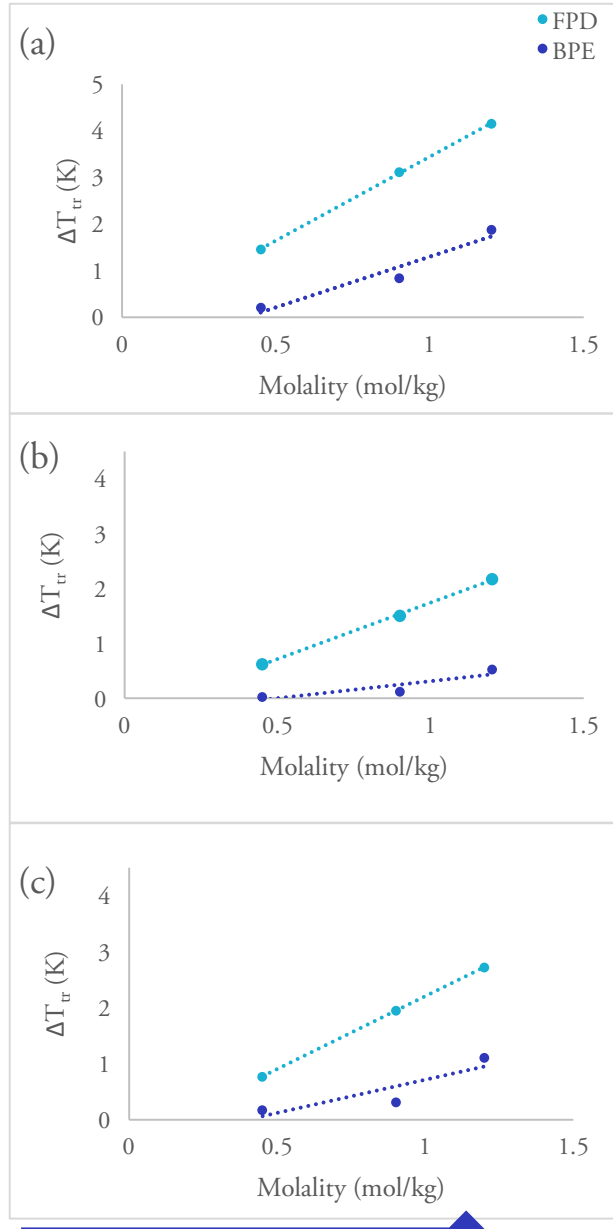


Figure 1: Variance of transition temperatures for (a) sodium chloride, (b) acetic acid, and (c) sucrose. FPD and BPE have been mapped on the same axes for comparison.

Table 1: Transition constants for sodium chloride, acetic acid, and sucrose.

	K_{fp} (°C/mol/kg)	K_{bp} (°C/mol/kg)	$(K_{bp}/K_{fp})/i$
Sodium chloride	3.6	2.2	0.3002
Acetic acid	2.1	0.63	0.3017
Sucrose	2.6	1.2	0.4536
			0.35 ± 0.09 s.d.

Qualitatively, there seemed to be no relationship between FPD and BPE as shown by the different slopes and by the absolute value of FPD being consistently higher than the absolute value of BPE. However, quantitatively, the rates at which these variance of transition temperatures change by concentration of solute are outlined in Table 1. Normalizing the proportionality by the van't Hoff factors of the solutes, it has been found that boiling point elevates 35% as quickly as freezing point depresses upon changing concentration.

DISCUSSION

The result of this study provides experimental evidence to support and demonstrate the relationship between FPD and BPE. An expression can be derived via the general relationship of colligative properties. In general, the rate by which the freezing point of a substance declines by the amount of dissolved solute is described as follows: $K_{fp} = \frac{M_A R T_m^{*\,2}}{\Delta_{\text{fus}} H_A}$ (Ge and Wang, 2009). Similarly, the rate by which the boiling point of a substance elevates by the amount of dissolved solute is described as follows: $K_{bp} = \frac{M_A R T_b^{*2}}{\Delta_{\text{vap}} H_A}$. Therefore, the proportionality of BPE against FPD can be determined simply by dividing the two expressions to arrive at the following expression:

$$\frac{K_{bp}}{K_{fp}} = \left(\frac{T_{bp,pure\ solvent}}{T_{fp,pure\ solvent}}\right)^2 \left(\frac{\Delta_{fus} H_{pure\ solvent}}{\Delta_{vap} H_{pure\ solvent}}\right).$$

This is a proportionality relating the variance of transition temperature in terms only of the pure solvent. In the case of water, as studied in this experiment, the boiling point and freezing point of water without any solute is 373.15 K and 273.15 K, respectively. It is also known that the enthalpy of fusion of water at atmospheric pressure and at the freezing temperature is 333.5 J/g and the enthalpy of vaporization of water at atmospheric pressure and at the boiling temperature is 2260 J/g. With this information, the theoretical value that describes the rate at which the BPE of water increases in proportion to its FPD is 0.2754. In other words, theoretically, the BPE of water should increase 27.54% as quickly as the decline of FPD.

The experiments in this study could be incorporated in the thermodynamics laboratory curricula to provide an interactive way to teach colligative properties in addition to allowing students to develop skills in planning experiments to answer specific research questions. This study only looked at water as the query solvent; however, a proportionality for different solvents such as diethyl ether could be calculated theoretically via the relationship expressed above and experimentally tested with the procedure provided in this study. Furthermore, this laboratory experiment has room for collaborative efforts in that groups of students could run this procedure for specific sets of conditions and all the data of every student could be collected and analyzed. The students can vary many parameters including: the solvent, the choice of solute, and the amount of solute dissolved. The students could also all run the same experiment and the same conditions and simply pool all the data to increase statistical significance and achieve more accurate results. The specific skills that the students will develop include: how to calculate molarity/molality, basic laboratory skills, and statistical analysis. Overall, the experimental procedure designed in this study provides a simple and reproducible method of describing the theoretical phenomena of FPD and BPE to be applied as a pedagogical tool to teach students about the relationship between two colligative properties.

ACKNOWLEDGEMENTS

Not enough acknowledgements can be given to Russ Ellis for his extremely supportive and constructive supervision throughout this research endeavour.

REFERENCES

Ge X, Wang X., 2009. Estimation of Freezing Point Depression, Boiling Point Elevation, and Vaporization Enthalpies of Electrolyte Solutions. *Industrial & Engineering Chemistry Research*, 48 (4), 2229–2235.

Human Papilloma Virus: Combatting Cancer with Vaccination

DAKOTA BINKLEY*, JACQUELINE WATT*, ACHINT BAJPAI†
Integrated Science* and Life Sciences† Programs, Class of 2017, McMaster University

SUMMARY

Human papilloma virus (HPV) is a virus responsible for cancer progression, and it has been directly associated with the development of cervical cancer. HPV-positive tumours have also been found in oral cavities, creating a rise in head and neck cancers. Recently, a vaccine has been developed that is highly effective at preventing the pathogenesis of HPV and HPV-related cancers. This article reviews the basic progression of HPV in the body and expands on the newfound dangers associated with this prevalent sexually transmitted disease.

Received: 03/05/2017 **Accepted:** 04/21/2017 **Published:** 04/30/2017
URL: https://journals.mcmaster.ca/iScientist/article/view/1455

Keywords: Human Papilloma Virus (HPV), vaccination, viral pathogenesis, cancer

INTRODUCTION

Human papilloma virus (HPV) is the most common chronic sexually transmitted disease (Workowiski and Berman, 2010). This can be attributed to HPV's ability to transmit through condoms, while producing few symptoms in the host (Watson, 2005). In severe cases, those infected with HPV can develop high-risk cancers, such as cervical cancer (Burd, 2003). Due to the disease's widespread prevalence, public health experts produced a vaccine that immunizes for two of the most prevalent high-risk (HR) strains of HPV; HPV-16 and HPV-18 (National Cancer Institute, 2016). This vaccine was originally advertised for preteen girls, who are less likely to be sexually active (National Cancer Institute, 2016) thus reducing their exposure to HPV and subsequent risk of cervical cancer. The most current vaccine covers nine of over a dozen HR strains of HPV (National Cancer Institute, 2016).

During the years 2007 to 2012, Canada only vaccinated young girls, as HR HPVs are directly associated with the development of cervical cancer (Public Health Agency of Canada, 2015). The recent decrease in cervical cancer has been attributed to the effectiveness of the vaccine (Brotherton, et al., 2011). Interestingly, a steady increase in head and neck cancers has been reported, especially in unvaccinated HPV positive individuals (Kimple and Harari, 2015). These tumours are prominent in males; however, women can also have HPV positive oral tumours. It is hypothesized that this increase of head and neck cancers is associated with modern sexual attitudes, including the increase of oral sex performed on partners (Marur, D'Souza, Westra and Forastiere, 2010). As both the vaginal tract and the oral cavity are lined by mucosal cells, for which HPV is tropic, the virus can be easily spread between the female genital tract and the oral cavity, providing optimal conditions for potential cancer development. This paper aims to present a comprehensive overview of HPV, including its virology, cancer pathogenesis, and the importance of vaccination.

HPV STRUCTURE

HPV is a small (50-60 nm diameter) virus that consists of a genome of 8000 base pairs (Figure 1) (Doorbar et al., 2015). This encodes for the viral proteins that are required for both viral spreading and invasion. The virus encodes for both capsid proteins and virulent proteins that are responsible for the structural integrity and pathogenesis of

the virus, respectively. The capsid proteins include L1 and L2. Meanwhile, the E2 protein acts as a regulator for the expression of viral proteins downstream of itself in the genome, particularly E6 and E7. The E6 and E7 proteins are essential for establishing an infection and immune evasion.

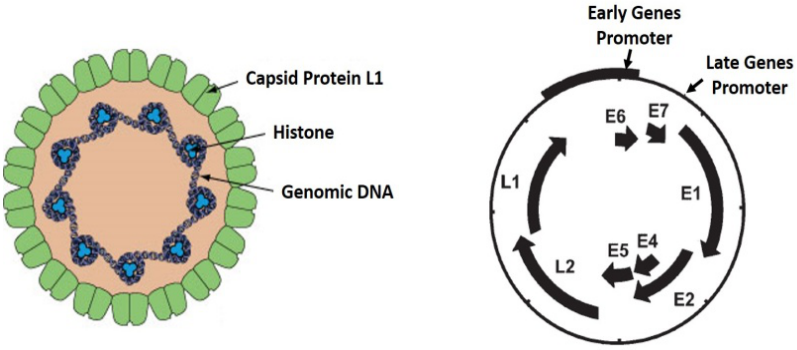


Figure 1: The structure of HPV. The image on the left shows the structure of the virion with the main capsid protein L1. The image on the right shows the genome organization of HPV. The most important proteins for viral pathogenesis are the early genes, E6 and E7; their regulator, E2; and the late capsid proteins, L1 and L2 (Hebner and Laimins, 2006).

INITIAL INFECTION

HPV can be classified as either low risk (LR) or HR, where LR HPVs result in non-cancerous genital warts and HR HPVs can result in either warts or gradually develop into cancer within the cervix or oral cavity (Münger, et al., 2004). This difference in pathogenesis is due to a decreased effectiveness of E6 and E7 in the LR strains (Münger et al., 2004). Both LR and HR HPVs progress identically though the mucosal layers (Figure 2). Infection initiates at the basal layer of the skin, and progresses to the spinosum layer (Longworth and Laimins, 2004). At this point, the virus enters neighbouring cells. The infection continues, passing through the upper layers of the skin – the granulosum layer and corneum layer. At this stage, uninfected cells would come out of the cell cycle to become keratin-filled sacs that make up the top layer of skin. However, the E7 protein of HPV is able to alter P21, a key cell cycle regulating protein, causing the infected cell to remain in the cell cycle. This ensures that the cellular replication machinery remains functional to continue to replicate the viral genome.

CHRONIC INFECTION

In most cases, the host is able to clear the virus using both the innate and adaptive immune pathways (Scott, et al., 2013). Clearance occurs within the first six months of being infected for 80-90% of individuals; however, during this time the virus can still be spread from the infected individual (Molano, et al., 2003). LR HPVs are more likely to be cleared in this time, with HR HPV infections taking up to one year to be cleared. However, if the virus successfully evades the host's immune system, it can eventually progress to a chronic infection for the remaining 10-20% of individuals. It is these individuals that risk developing cancer (Burd, 2003).

Once a chronic infection is established, one of the main factors that determines the development of cancer is the random integration of the viral genome into the host's genome (Williams, Filippova, Soto and Duerksen-Hughes, 2011). The viral genetic information is typically integrated in common fragile sites of the host's chromosomes, which are found in every individual (Thorland, Myers, Gostout and Smith, 2003; Williams, et al., 2011). When integration occurs, the viral genome is cleaved at the E2 gene. The downstream E6 and E7 genes are then integrated into the host. Since E2 is a regulator of E6 and E7, their expression is no longer controlled in the host chromosome, resulting in their overexpression. These proteins, particularly E7, are responsible for keeping the cell in its replicative cycle. When they are overexpressed, this can lead to uncontrolled cell growth and cell immortalization, which are hallmarks of cancer.

Integration into the host genome does not guarantee the development of cancer as the host can still combat the effects of chromosome integration. Those who develop cancer typically have environmental co-factors that make the cell more likely to become cancerous (Wang, et al., 2009). These co-factors include smoking, drinking, and hormonal birth control use. Each

of these factors can directly change the mucosal cell environment in either the cervix or the oral cavity.

HPV-SPECIFIC CANCER PROGRESSION

In an HPV transfection study on human epithelial cells, E6 and E7 were shown to be directly involved in the transformation of these cells into malignant cancers. Hawley-Nelson, et al. (1989) created a plasmid consisting of the genome segments for E6 and E7. They placed transcription termination sequences at E6, E7, or both segments to prevent the translation of these proteins. When E7 was inhibited, no proliferative cell activity was observed. When E6 was inhibited, there was some proliferation, however at a slower rate than the positive control group. In addition, senescence was observed after two months of observation. When both E6 and E7 were permitted for transcription, proliferation continued for several months with no sign of senescence. This experiment supported the importance of E6, and particularly E7, in transforming infected cells (Hawley-Nelson, et al., 1989).

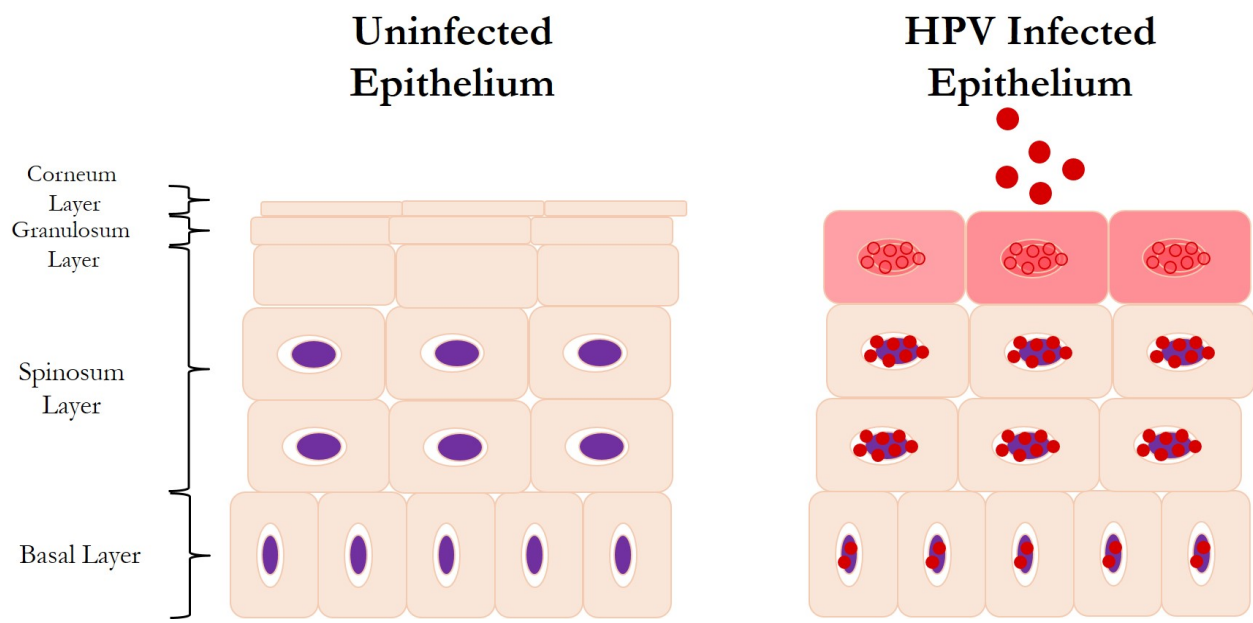
E6 and E7 are critical to the cancer progression associated with HPV integration as both can create complexes within cells that inhibit natural cell cycle regulation checkpoints (Münger, et al., 2004). For instance, E6-mediated viral pathways can result in the degradation of p53, a protein involved in cell cycle checkpoints (Bernard, et al., 2011; Münger, et al., 2004). E6 can also create binding complexes that allow for HPV-infected daughter cells to have never-ending telomerase, a genome sequence that signals for cell death when it is used up, resulting in the evasion of apoptosis (Oh, Kyo and Laimins, 2001).

E7 is arguably more pathogenic than E6, as it typically increases the rate of E6 reactions and initiates its own cellular pathways (Yim and Park, 2005; Münger, et al., 2004). E7 can inhibit the function of the Retinoblastoma protein (pRB), allowing cells to enter the S1 phase of the cell cycle without regulation, bypassing a major cell cycle checkpoint. The manipulation of these pathways by viral E6 and E7 results in uncontrolled proliferation of HPV infected cells, resulting in tumour formation.

PUBLIC HEALTH CONSIDERATIONS

In 2007, the annual financial burden of HPV was $300 million, with nearly 18% of this cost coming from the treatment of genital warts and cervical cancer (Lalonde, 2007). The quadrivalent recombinant vaccine covering HPV 6, 11, 16,

Figure 1: Comparison between uninfected and infected epithelium upon initial infection by HPV. HPV enters the basal layer through microabrasions. From here it establishes a low resident viral load that is passed on when the basal cell divides. HPV then moves through the layers of the skin as the epithelial cells continue to divide. Figure modified from (Longworth and Laimins, 2004).



and 18 is a strong solution to reduce the financial burden while improving public health. A fully efficacious vaccine administered to 12-year-old Canadian girls would prevent nearly 10,000 cases of genital warts and 140 cases of cervical cancer per 100,000 vaccinations. In addition, the estimated disease-associated cost savings would amount to $7.7 million per 100,000 vaccinations, which can greatly influence healthcare spending towards other, non-preventable illnesses (Lalonde, 2007). Recently, these vaccines have become available to males as well in hopes of decreasing the current rates of HPV positive oral cancers. It is critical to vaccinate for HPV, as vaccination costs less than cancer treatment and because it can prevent future suffering of HPV-positive cancers, whether they are cervical or head and neck related.

CONCLUSIONS

Due to the obvious effectiveness of the HPV vaccine, it is recommended by Health Canada that vaccination of non-sexually active pre-teen females and males be continued across the nation. As HPV-related cancers are attributed to great health and financial burdens, preventing the spread and development of HR HPVs is crucial to the population. Now that researchers understand the pathogenicity of the HPV virus, more effective vaccines can be produced that allow for the at-risk population to be immunized against HR and potentially LR HPV types.

REFERENCES

Bernard, X., Robinson, P., Nominé, Y., Masson, M., and Brutscher, B., 2011. Proteasomal Degradation of p53 by Human Papillomavirus E6 Oncoprotein Relies on the Structural Integrity of p53 Core Domain. *PLoS ONE*, 6(10), p.e25981.

Brotherton, J.M., Fridman, M., May, C.L., Chappell, G., Saville, A.M., and Gertig, D.M., 2011. Early effect of the HPV vaccination programme on cervical abnormalities in Victoria, Australia: an ecological study. *The Lancet*, 377(9783), pp.2085-2092.

Burd, E.M., 2003. Human papillomavirus and cervical cancer. *Clinical microbiology reviews*, 16(1), pp.1-17

Doorbar, J., Egawa, N., Griffin, H., Kranjec, C., and Murakami, I., 2015. Human papillomavirus molecular biology and disease association. *Reviews in Medical Virology*, 25(S1), pp.2-23.

Hawley-Nelson, P., Vousden, K., Hubbert, N., Lowy, D., and Schiller, J., 1989. HPV 16: E6 and E7 protiens cooperate to immortalize human foreskin keratinocytes. *EMBO Journal*, 8(12), pp.3905-3910.

Hebner, C.M., and Laimins, L.A., 2006. Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity. *Reviews in Medical Virology*, 16(2), pp.83-97.

Kimple, R.J., and Harari, P.M., 2015. The prognostic value of HPV in head and neck cancer patients undergoing postoperative chemoradiotherapy. *Annals of translational medicine*, 3(Suppl 1), p.S14.

Longworth, M.S., and Laimins, L.A., 2004. Pathogenesis of Human Papillomaviruses in Differentiating Epithelia. *Microbiology and Molecular Biology Reviews*, 68(2), pp.362-372.

Marur, S., D'Souza, G., Westra, W.H., and Forastiere, A.A., 2010. HPV-associated head and neck cancer: a virus-related cancer epidemic. *The Lancet Oncology*, 11(8), pp.781-789.

Molano, M., van den Brule, A., Plummer, M., Weiderpass, E., Posso, H., Arslan, A., Meijer, C.J.L.M., Muñoz, N., and Franceschi, S., 2003. Determinants of Clearance of Human Papillomavirus Infections in Colombian Women with Normal Cytology: A Population-based, 5-Year Follow-up Study. *American Journal of Epidemiology*, 158(5), pp.486-494

Münger, K., Baldwin, A., Edwards, K.M., Hayakawa, H., Nguyen, C.L., Owens, M., Grace, M., and Huh, K., 2004. Mechanisms of human papillomavirus-induced oncogenesis. *Journal of virology*, 78(21), pp.11451-60.

National Cancer Institute, 2016. *Human papillomavirus (HPV) vaccines*. U.S Department of Health and Human Services.

Oh, S.T., Kyo, S., and Laimins, L.A., 2001. Telomerase activation by human papillomavirus type 16 E6 protein: induction of human telomerase reverse transcriptase expression through Myc and GC-rich Sp1 binding sites. *Journal of virology*, 75(12), pp.5559-66.

Public Health Agency of Canada, 2015. *Human papillomavirus (HPV) prevention and HPV vaccines: questions and answers*.

Scott, M.E., Shvetsov, Y.B., Thompson, P.J., Hernandez, B.Y., Zhu, X., Wilkens, L.R., Killeen, J., Vo, D.D., Moscicki, A.-B., and Goodman, M.T., 2013. Cervical cytokines and clearance of incident human papillomavirus infection: Hawaii HPV cohort study. *International journal of cancer*, 133(5), pp.1187-96.

Thorland, E.C., Myers, S.L., Gostout, B.S., and Smith, D.I., 2003. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene*, 22(8), pp.1225-1237.

Wang, S.S., Zuna, R.E., Wentzensen, N., Dunn, S.T., Sherman, M.E., Gold, M.A., Schiffman, M., Wacholder, S., Allen, R.A., Block, I., Downing, K., Jeronimo, J., Carreon, J.D., Safaeian, M., Brown, D., and Walker, J.L., 2009. Human papillomavirus cofactors by disease progression and human papillomavirus types in the study to understand cervical cancer early endpoints and determinants. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 18(1), pp.113-20.

Watson, R.A., 2005. Human Papillomavirus: Confronting the Epidemic-A Urologist's Perspective. *Reviews in urology*, 7(3), pp.135-44. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16985824> [Accessed 26 Feb. 2017].

Williams, V.M., Filippova, M., Soto, U., and Duerksen-Hughes, P.J., 2011. HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. *Future virology*, 6(1), pp.45-57.

Workowiski, K., and Berman, S., 2010. *Sexually Transmitted Diseases Treatment Guidelines, 2010*.

Yim, E.-K., and Park, J.-S., 2005. The role of HPV E6 and E7 oncoproteins in HPV-associated cervical carcinogenesis. *Cancer research and treatment : official journal of Korean Cancer Association*, 37(6), pp.319-24.

THANK YOU

The Editorial Board of *The iScientist* would like to extend our sincerest thanks to all of the individuals and organizations who helped support our journal from its inception to the publication of this second volume.

Of special note, we want to thank the administrators of the Science Initiative Fund of the McMaster Science Society. Their financial support and belief in this project has afforded us the opportunity to provide print copies of this journal for the next few years. In addition to this, they have graciously provided support for a series of workshops regarding undergraduate publishing and the peer-review system that will be hosted by the Editorial Board and made available to all interested students at McMaster University.

In no particular order, we would also like to thank:

Dr. Sarah Symons and Andrew Colgoni, Editors-in-Chief of this journal, for their logistical support and continued advising over the course of this journal's history

The Integrated Science Faculty for their support of the journal and their perseverance in teaching the fundamental philosophies on which it is built

The Integrated Science Student Body for their contributions to the journal and the patience and respect they showed to the peer-review process and our Board

The School of Interdisciplinary Science for their recognition of the value of multidisciplinary research and contributions to the journal through their students

McMaster University Publishing System for Student Journals for hosting our website and **Open Journal Systems** for providing a robust software for receiving articles and disseminating feedback

Friends and Family for their continued support amidst long nights of editing and formatting

CONTACT US

journals.mcmaster.ca/iScientist

iscijournal@gmail.com

H.G. Thode Library. Room 306
1280 Main Street West
Hamilton, Ontario L8S 4P5

COVER PHOTO

A *Meloe angusticollis*, or a short-winged blister beetle, found in the McMaster Forest Teaching and Research Facility

Photograph by Kyra Simone, 2017

ISSN 2560-9947

THE iSCIENTIST