MURJ

Massachusetts Institute of Technology Undergraduate Research Journal

Iceland

Qatar

Singapore

MURJ

Massachusetts Institute of Technology
Undergraduate Research Journal

Contents

Cover image: Modified figure from "Gathering Real World Data to Predict Admission Statistics;" See page 28 for the full report

May 2015

Dear MIT Community,

I am delighted and honored to introduce the Spring 2015 issue of the MIT Undergraduate Research Journal (MURJ) to you. This particular issue marks the 15th year since MURJ was founded. Despite the passage of time, the original vision for the publication—an interdisciplinary, peer-reviewed scientific journal, modeled after Scientific American—still holds true today in this edition of the journal.

MURJ offers a window into the remarkable research students undertake here through the Undergraduate Research Opportunities Program (UROP). UROP was established at MIT in 1969 and is now emulated at academic institutions around the globe. In many ways, UROP embodies what MIT's founder, William Barton Rogers, envisioned for his then-novel, independent educational institution: learning by doing, and bringing that knowledge to bear on the nation's—and the world's—greatest challenges. Since its inception, UROP has become an integral part of an MIT education: in any given year, 60% of undergraduates participate. And by the time they graduate, 89% of students have completed at least one UROP. These are remarkable statistics that Rogers would undoubtedly be pleased with!

Participating in faculty-mentored research is a powerful experience for undergraduates. However, the same can be said of the faculty experience. I have been privileged to collaborate with a number of UROP students over the years, and it has been one of the most rewarding aspects of my work here.

MURJ also reflects the incredible breadth of research that takes place on a daily basis at the Institute, and the interdisciplinary nature of that work. Reaching across academic departments and school boundaries is an integral part of MIT; it's no accident that there are over 50 interdisciplinary centers, labs, and programs here, including the new Institute for Data, Systems and Society set to launch in July.

I hope you enjoy this issue of MURJ as much as I have. It is a wonderful distillation of just a fraction of the extraordinary and unique undergraduate research projects going on in all corners of the Institute.

Sincerely,

Denny Freeman

Dean for Undergraduate Education
Professor of Electrical Engineering
MacVicar Faculty Fellow

May 2015

Dear MIT community,

We hope you enjoy 29th issue of the MIT Undergraduate Research Journal, a biannual student-run publication that features groundbreaking undergraduate research from across campus. *MURJ* is now in its 15th year of publication, and the research conducted and articles written by MIT undergraduates demonstrate, as always, phenomenal passion and talent in all scientific disciplines. The high diversity of research interests among MIT undergraduates is evident--in this issue, we learn about a computational model predicting admissions statistics of MIT international students, a method for improving the carbon-fixation abilities of a microorganism, and a comprehensive analysis of mortgage refinancing options.

In addition to our research articles, we also present science news articles reviewing cutting-edge scientific discoveries, as well as features articles elaborating more comprehensively on particularly fascinating topics. In this issue, our features detail the creativity of science education and outreach to the general public, through the merging of artistic and scientific work in mediums such as fashion and video games. We also interview and spotlight Professor David C. Page, Director of the Whitehead Institute, to learn his path through science and his advice to our peer young scientists at MIT.

During CPW 2015, *MURJ* hosted its second interdisciplinary undergraduate poster session, currently the only annual all-major poster session symposium at MIT. The event was a great sucess, and we are pleased to include an article reviewing the event. We invite you to check out this article and the others presented within on our website at murj.mit.edu, where we highlight

the work of the MIT students published within and host an archive of previous *MURJ* issues.

This journal is a collaborative effort by an extraordinary team of dedicated students, and we would like to thank all our editorial board and contributors for their time and effort this semester. Without the work of our hard-working staff members, *MURJ* would not exist. We would also like to thank all of the student researchers who graciously shared their research.

If you would like to contribute to future issues of the MIT Undergraduate Research Journal, we invite you to join our team of authors and editors or submit your research for our Fall 2015 issue. Please contact murj-officers@mit.edu if you have any questions or comments.

Best,

Elliot Akama-Garren
(Co-Editor-in-Chief)

Tatyana Gubin
(Co-Editor-in-Chief)

Reuben Saunders
(Co-Editor-in-Chief)

Science News In Review

Tattoos: Glucometers of the 21st Century

Long gone may be the days where diabetes patients have to puncture themselves with needles in order to measure their blood sugar levels. In a paper published in Analytical Chemistry, scientists at the University of California, San Diego have presented a proof-of-concept of a temporary tattoo that has the ability to monitor glucose levels in a noninvasive manner.

Blood glucose measuring tattoo
Credit: http://news.sciencemag.org/chemistry/2015/01/temporary-tattoo-measures-blood-glucose-levels

This tattoo uses iontophoretic-biosensing to measure glucose levels and has been shown to be selective for glucose and to be able to detect rapid rises in glucose levels after eating. Currently, the tattoo only works for a day, and it does not give a numerical read-out of blood sugar content; the scientists are now working on adding this feature while also increasing the tattoo's longevity and adding new functions such Bluetooth connectivity, which would allow the diabetes patients' tattoos to send collected data to their doctor. Nevertheless, this tattoo represents a monumental advancement

in the development of a non-invasive glucose monitor.

—S. Santiago

Source: http://news.sciencemag.org/chemistry/2015/01/temporary-tattoo-measures-blood-glucose-levels

Matter-Antimatter Asymmetry Link to Higgs Boson

The planets and stars that we see around us are made up of matter. This ordinary matter has a counterpart – antimatter. Particles of antimatter are identical in size and mass to that of ordinary matter, but have opposite other properties such as charge and spin. A notable effect of interaction between matter and antimatter is annihilation where the particle and antiparticle collide and most often disappear emitting energy.

For several decades now, cosmologists and astronomers alike have been perplexed by the fact that our observable universe is abundantly more matter than antimatter, an imbalance known in cosmology as asymmetry. Alexander Kusenko and his colleagues at UCLA believe that in the early universe the Higgs field may have influenced the proportion. The Higgs field is associated with the Higgs boson, the particle recently discovered by physicists at CERN working at the LHC, and is responsible for giving particles and antiparticles their mass.

In the primordial soup of the early universe, there existed many more particles than we see today. It is believed that antimatter and matter existed in nearly equal proportions with a difference of only one in 10 billion. As the universe cooled, they annihilated to leave just the matter that we see today. Kusenko believes that this slight difference was the result of a process called Higgs relaxation wherein the Higgs field changes with the course of the universe's evolution. Researchers at the LHC have measured this effect through the mass of the Higgs boson and found the field to be more representative of the first

All the planets and stars in the observable universe are made up of matter and not antimatter. This imbalance is what is known as baryon asymmetry.
Credit: NASA

Mapping the Success of Startups

Entrepreneurial Quality
Percentile Distribution

Guzman & Stern, Science (2015)
Illustration by RJ Andrews

Map portraying the entrepreneurial quality in California's Bay Area. Notable concentration on the South Bay centered on Stanford University. .

Credit: GUZMAN AND STERN/SCIENCE (2015); (ILLUSTRATION) R. J. ANDREWS

in California and found Silicon Valley as the most dominant entrepreneurial hotspot beating out San Francisco, San Diego and Los Angeles. More specifically, they found higher entrepreneurial quality located around research institutions, and in particular Stanford, which lies at the center of Silicon Valley. The researchers believe the work has important implications on entrepreneurial growth dynamics, scientific research, and possibly public policy.

—P. Nagaraj

Sources: http://www.sciencemag.org/cotent/347/6222/606.full.pdf
http://news.sciencemag.org/economics/2015/02/why-facebook-and-google-succeeded

A basic checklist of key factors for success of a startup includes: location, trademarking, patents, legal incorporation, and naming.

MIT affiliates Jorge Guzman and Scott Stern tracked various entrepreneurial ventures across California and reinforced some common perceptions as well as elucidated some new factors that lead to the success of a startup. For instance, having a shorter name like Google or Facebook proves to be more effective than longer company names such as Cryptine Networks.

Establishing a relationship between entrepreneurship success and regional location has been difficult. Much of the previous attempts focused on quantity of ventures, whereas in the study by Guzman

and Stern they attempt to assess the quality of ventures in order to overcome this difficulty. They introduce a method to measure what is called the entrepreneurial quality index. By accounting for naming decisions, type of business registration, and control of intellectual property, the researchers were able to predict the likelihood of company growth and in turn compare the entrepreneurial quality across cities in the state of California.

Guzman and Stern found that names associated with high-technology clusters are indicative of growth whereas companies named after the founders are less likely to grow. Registering as a corporation and trademarking also suggest a higher probability of growth in a company. They then mapped this quality index against locations

few moments of the Big Bang than what we expect today. Therefore as a result, Kusenko states that particles and antiparticles may have had temporarily unequal masses due to the early Higgs field leading to the asymmetry we have today.

—P. Nagaraj

Source: http://www.sciencedaily.com/releases/2015/02/150225132255.htm

COGNITIVE SCIENCE

False Memories in Bumblebees

Bumblebees, like humans, can make false memories.
Credit: http://www.sciencedaily.com/releases/2015/02/150226131940.htm

In a recent study that was among the first to explore false memories in non-human animals, scientists were able to ascertain that bumblebees are prone, much like humans, to remember things incorrectly. To conduct their research, Lars Chitika and his colleague, Katheryn Hunt of Queen Mary University of London trained bumblebees to expect a reward when visiting a solid yellow artificial flower followed by one with black and white rings or vice versa. During test trials, the bees were given a total of three choices among flowers, the third of which was a mixed-up version of the other two. For example, the third flower would have yellow and white rings. If this test was conducted only minutes after their training, the bees would gravitate towards the one that they had most recently been rewarded with – their short term memory was therefore intact. After a period of one to three days, however, something very different happened when the bumblebee's memory was put to the test: They began to select the flowers that had yellow rings, even though that was not a flower they had used in their training before.

Chitika and Hunt state that the merging of long-term memories phenomenon observed in these bumblebees is what happens to humans as well. They believe that these false memories are not "bugs in the system" but rather the side effects of an adaptive learning memory system that is parallel to human memory formation. Further research is being conducted and, so far, the group has been able to find that extracting patterns and commonalities between different events in our environment is an inherent part of our ability to adapt to our environment. Thus, the integration of lifetime experiences can accumulate and be incorporated into a comprehensive foraging decision as in the case for bees and life decisions as in the case of humans.

—K. Ziadkhanpour

Source: http://www.sciencedaily.com/releases/2015/02/150226131940.htm

A tornado in Shoal Creek Valley, Alabama, one of the 120 to hit southeastern United States on April 27, 2011
Credit:http://upload.wikimedia.org/wikipedia/commons/9/9c/Shoal_Creek_Valley_Alabama_Tornado_April_27%2C_2011.jpg

EARTH SCIENCE

Smoke Found to Increase Tornadoes' Strength

Setting fires to clear farmland is an agricultural technique around the world, but recent research suggests that the smoke from these fires may lead to more violent hurricanes. A paper recently published in Geophysical Research Letters looks at the deadly outbreak of 120 tornadoes that hit the southeastern United States in the Spring of 2011. Researchers noted that the air in the region contained smoke that could be traced back to Mexican and Central American agricultural fires set earlier that year. Using computer models to recreate the atmospheric conditions at that time, researchers believe that the smoke in the air could have bolstered the tornadoes' force. In particular, the model produced faster wind speeds and larger storm clouds than the smoke-free simulation. These weather conditions- the fast wind and low, thick clouds- would not be enough to have produced the tornadoes on their own, but their presence did make these storms more intense. Perhaps the tornadoes would not have racked up such a large death toll, ultimately killing 313 people, if

A photograph of the Atmospheric Radiation Measurement Climate Research Facility in Hawaii, one of two Department of Energy sites where thermal radiation measurements were recorded for ten years.

Credit: http://commons.wiki-media.org/wiki/File:ARM_facility_Nauru.jpg

there had not been smoke in the air to boost their strength. At the very least, these researchers hope that their results will encourage meteorologists to factor smoke into their future weather models.

—E. Berg

Source: http://news.sciencemag.org/earth/2015/02/smoke-distant-fires-could-create-more-deadly-tornadoes

Experimental Confirmation that CO₂ Contributes to Global Warming

Carbon dioxide is an important greenhouse gas. Atmospheric carbon dioxide traps heat that reflects of the Earth's surface, preventing this heat from escaping the atmosphere. Over time, this increase in trapped energy raises temperatures around the world. While this has been common knowledge for years, scientists have only recently confirmed this theory experimentally, finding an increased amount of thermal radiation absorbed at two data collection sites over the last ten years. These researchers, working at the Department of Energy's Lawrence Berkeley National Laboratory, had been collecting almost-daily radiation measurements at research facilities in Oklahoma and Alaska from 2000 to 2010. To make these measurements, the researchers

used spectroscopic instruments that measured both the total amount of infrared energy trapped in the atmosphere and the radiation specifically emitted by carbon dioxide. This data shows a 0.2 Watt per square meter increase in thermal energy in both locations over this time period. This increase is consistent with what theory would predict for the amount of carbon dioxide that has been emitted into the atmosphere over this same time period. This research is the first definitive and experimental proof that global carbon dioxide emissions are warming the surface of the Earth.

—E. Berg

Source: http://www.sciencedaily.com/releases/2015/02/150225132103.htm

GENETICS

The Darker History Hidden in Our Genes

Today we know that the rich nuances hidden in genomic data speak volumes about people and their ancestors. But, in digging deep back into the genetic variations that trace to our early evolutionary history, could we possibly unearth the origin of uniquely human disorders? That is could the same insights from genomic data that have previously shed light

on human evolution also speak to our understanding of mental illnesses like autism, schizophrenia, and bipolar disorder? A recent study published by Mount Sinai researchers in Molecular Biology and Evolution may have the answer.

The team of researchers led by Joel Dudley chose to reexamine data from the Psychiatric Genomics Consortium- a compendium that boasts genomic data from 36,989 schizophrenia cases and 113,075 controls. Specifically, the group sought to explore a possible relation between previously identified schizophrenia-associated loci and human accelerated regions (HARs). HARs represent hotspots of high mutation activity in the human genome relative to non-human species; in short, these small stretches in the human genome in tandem with nearby genes conserved in non-human primates (pHARs) may very well be the crux of what separates humans from all other species.

Bayesian network of genes related to prefrontal cortex development. Non HAR-associated schizophrenia genes (blue), HAR-associated schizophrenia genes (pink), other genes (green).

Credit: http://mbe.oxfordjournals.org/content/early/2015/03/15/molbev.msv031.full.pdf+html

What Dudley and his colleagues discerned was that pHARs are heavily enriched with schizophrenia-associated loci. Moreover, they discovered that the schizophrenia genes associated with pHARs seem to be under a stronger selective pressure than other schizophrenia genes. In the short term, their analysis points to a prioritization of genomic targets for treating schizophrenia; knowing that pHAR-associated schizophrenia genes are more likely to persist in the gene pool speaks to them being lucrative targets going forward. In the long term, similar evolutionary history based analyses could be applied for other neuropsychiatric disorders like autism and bipolar disorder. But for now it remains humbling that, within the complex genetic underpinnings of all that makes us human, lie lurking the crippling seeds of mental illness.

—A. Bhandarkar

COMPUTATIONAL BIOLOGY

DARPA Sets Out to Automate Cancer Research by 2017

The U.S. Defense Advanced Research Projects Agency (DARPA) has set out to automate cancer research by 2017 with the launch of the $45 million program Big Mechanism last summer.

Big Mechanism aims to develop computer systems that will integrate information from research papers and computer models of cancer mechanisms in order to frame new hypotheses for cancer development. According to the program's manager Paul Cohen, its goal is to help scientists deal with the complexity of cancer in a time where research has become highly specialized. If Big Mechanism succeeds, its methodology could be applied to complicated systems in fields ranging from climate science to military operations to poverty. Currently, though, the program's focus is on untangling complex cancer pathways caused by mutations in the Ras gene family, which account for about one third of human cancers. To do so, DARPA is developing systems that will read the literature on Ras-driven pathways, identify and extract useful information, and convert this information into formal representations that can then be integrated into models of Ras pathways. As of January, Big Mechanism's best-performing machine reading system extracted 40% of useful information from six paragraphs of text and correctly determined how it related to rudimentary models of Ras-driven cancer pathways. According to Cohen, this is an excellent start. He is confident that, one way or another, Big Mechanism will pay off. "DARPA seeks revolutionary technology," he said. "Sometimes those technologies are turned into practice.

DARPA's Big Mechanism project will use computers to investigate the mechanisms responsible for the formation of cancerous cells, such as those depicted in the stylized image above.

Credit: Image courtesy of dream designs at FreeDigitalPhotos.net

Sometimes, they show the world what is possible."

—J. Switzer

ANTHROPOLOGY

Morality-Based Religions May Have Their Roots in Wealth

An increase in wealth may have caused the transition from ritual- to morality-based religions, proposes a new study from the École Normale Supérieure (ENS) in Paris. Nicolas Baumard, a psychologist at ENS, explains that it wasn't until around 500 B.C.E. that new religions focusing on morality, self-discipline, and asceticism appeared in Eurasia. Before that time, religion was largely focused on rituals and short-term rewards. Using historical and archaeological data, Baumard and his team have shown that this transition— also known as the 'Axial Age'— coincided with a rise in affluence throughout Eurasia. As people became increasingly wealthy, their comfortable living allowed them to shift their focus from short term goals—for example a good growing season—to longer term considerations such as what would happen to them after death. Not coincidentally, said Baumard, the new religions that began to appear during the Axial Age— Stoicism, Jainism, and Buddhism— and their successors, Christianity and Islam, promoted self-discipline and short-term sacrifice while promising a pleasant afterlife for the faithful. Using historical and archaeological data from across Eurasia during the Axial Age, Baumard and his colleagues found that the best predictor of the emergence of one of these "moralizing religions" was an affluence index called "energy capture": the number of calories worth of food, fuel, and other resources available per day to each person in a given society. Societies that had an energy capture of more than 20,000 kilocalories per day had a much higher likelihood of seeing the emergence of a moralizing religion; those below 20,000

kilocalories almost never did. Said Baumard, "You need to have more in order to be able to want to have less."

—J. Switzer

Source: http://news.sciencemag.org/biology/2014/12/wealth-may-have-driven-rise-todays-religions

MICROBIOLOGY

Ebola: Small Molecule, Big Consequences

Electron micrograph of an Ebola virion.
Credit: http://www.cdc.gov/vhf/ebola/modules/flexslider/about-ebola.jpg

The Ebola virus is just one of many lethal viruses that are currently lurking and causing an alarming number of deaths in Africa. Scientists have been actively researching for effective ways to treat these viruses, with the goal of potentially creating vaccines that could eradicate them once and for all. One particular route that has proven to be attractive is the use of monoclonal antibodies to target these viruses. This possible therapeutic treatment was the one explored by two scientists at the Scripps Research Institute, who, after a six-year effort, managed to identify and uncover the structure of an antibody that binds to the Marburg virus, Ebola's cousin. By using a blood sample of a Marburg fever survivor, the two scientists studied how the antibody attached itself to the virus, discovering that its attachment prevented the virus from binding to a receptor and eventually its replication. Further studies then showed that this antibody also had the ability to bind to the Ebola virus, yet -interestingly- it was not present in the blood of Ebola fever survivors. These findings provide a template for potential new treatments that could be used combat the Marburg and Ebola fever. Furthermore, the cross reactivity of this antibody may allow it to be modified in such a way that would permits it use in the treatment of other viruses that share similarities to the Ebola and Marburg virus, giving us a tool to fight and prevent possible future outbreaks.

—S. Santiago

Source: http://www.sciencedaily.com/releases/2015/02/150226122328.htm

Starving Bacteria Share Nutrients

Nutrient-starved bacteria can latch onto other bacterial species in order to share their food, recent studies have suggested. Marie-Thérèse Giudici-Orticoni of Aix-Marseille University, France and her colleagues have found that when Clostridium acetobutylicum, which metabolizes glucose for growth, and Desulfovibrio vulgaris, which metabolizes lactate and sulfate, are cultured together in a glucose-only medium, the D. vulgaris survives by latching itself onto the C. acetobutylicum and sharing the latter's cytoplasm and proteins. In a separate study, Christian Kost of the Max Planck Institute for Chemical Ecology in Jena, Germany and his colleagues performed a similar experiment with Escherichia coli and Acinetobacter baylyi. The bacteria were mutated so that they could not produce certain essential amino acids and then cultured in a medium lacking the amino acid required by E. coli. Similar to the observations of the Aix-Marseilles' team, the E. coli was found to form nanotubes of up to 14 micrometers in order to connect to the A. balyli. The two bacterial species then exchanged cytoplasm and amino acids so that each received the amino acid it was lacking. According to Kost and his team, the results suggest that these

A stylized image of an E. coli bacterium, one of the bacterial species that was observed by scientists to share food with other species when nutrient deprived.
Credit: Image courtesy of Victor Habbick at FreeDigitalPhotos.net

bacteria function as interconnected entities rather than individuals.

—J. Switzer

Source: http://news.sciencemag.org/biology/2014/12/wealth-may-have-driven-rise-todays-religions

ASTROBIOLOGY

Unique Biochemistry May Be Concealing Alien Life

An entire biosphere of organisms with a chemical makeup distinct from what we see on earth may be going undetected by scientists, said Victoria Orphan of the California Institute of Technology in Pasadena at the annual meeting of the American Association for the Advancement of Science (AAAS). When searching for alien life, scientists typically look for features that indicate an earth-like planet, such as an abundance of liquid water. However, this search precludes the existence of life forms with an unfamiliar biochemistry— for instance, organisms that formed in liquid hydrocarbons rather than in water.

The possible existence of this overlooked "shadow biosphere" has some scientists stressing the need to enlarge the search for life to include indicators more general than those with which we are already familiar.

Scientists have found even the earth's most extreme environments, such as this Yellowstone hot spring, to contain life. Life on other planets however, may prove to be more elusive.

Credit: http://news.sciencemag.org/biology/2015/02/shadow-biosphere-might-be-hiding-strange-life-right-under-our-noses

One such "universal indicator" is a system that is out of equilibrium, says Carolyn Porco of the Space Science Institute in Boulder Colorado. Because life takes in and uses energy, it alters its environment. For instance, the earth would not have such an oxygen-rich atmosphere if it weren't for the actions of photosynthetic organisms. Porco has called on other scientists to come up with a working definition of life in order to give planetary scientists more guidance on what to look for. Such

a criteria would not only be useful for alien life, however; even here on earth, there may be unique life forms going undetected by scientists. Tucked away in the microscopic world are the earth's most diverse organisms, some of which have already called into question our understanding of life. Although viruses are currently classified as nonliving, recent discoveries of giant amoeba-infecting viruses with complex genomes blur the lines between living and non-living. Such unexpected discoveries, according to Orphan, suggest that we shouldn't allow our search for life be constrained by what we already know is out there.

—J. Switzer

Source: http://news.sciencemag.org/biology/2015/02/shadow-biosphere-might-be-hiding-strange-life-right-under-our-noses

Astrobiology: A New Hope

Life on Earth is inextricably linked to the presence of one molecule: water. As the "universal solvent", water is central in enabling the chemical reactions necessary for biological function. In short, life as we know it cannot exist without water.

However, recently, a team at Cornell University has modeled a new type of methane-based life form that could exist on Saturn's moon Titan. With surface temperatures reaching 292 degrees below zero Fahrenheit, Titan is covered in oceans not of water, but of hydrocarbons. In these cryogenic, organic oceans, cells could be composed of membranes made up of small organic nitrogen compounds--which the researchers have dubbed "azotosomes" ("azote" being the French word for nitrogen)--rather than phospholipids. To the surprise of the researchers, these azotosomes show the same sort of stability and flexibility in conditions found on Titan that phospholipid membranes show here on Earth. Using a molecular dynamics method

Cells composed of "azotosome" membranes may exist in Titan's hydrocarbon oceans.

Credit: http://www.nasa.gov/mission_pages/cassini/multimedia/pia11001.html`

that screened for methane-based compounds able to self assemble into membranes, the team identified acrylonitrile as the most promising candidate to form the building blocks of azotosomes. Acrylonitrile is found in Titan's atmosphere, and showed a strong barrier to decomposition. With a strong candidate to form cell membranes identified, the

next steps will be to study the possible analogs to reproduction and metabolism in this methane-based life form. The team hopes to ultimately send a probe to Titan itself and sample its oceans.

—C. Gonzalez

Source: http://www.sciencedaily.com/releases/2015/02/150227181333.htm

MURJ
Features

Science in the World: Fashion and Video Games

Science outreach is important for education and for invoking enthusiasm for research. Dr. Tal Danino is helping involve the public in biology research by using fashion and video games.

BY ETA ATOLIA

2014 Fashion Descience Runway Show - MIT Media Lab.
Photo Credit: Sharon Lacey

People associate many words with biological science: test tubes, gloves, lab coats, chemicals, bacteria, viruses, medicine, etc. In this list, 'fashion' and 'games' are nowhere to be found.

Current biology research is very sophisticated and very specific to relevant scientific topics. This allows for great research but inevitably alienates those without higher learning in the field. People throughout MIT and the science community are trying to amend these gaps via different outreach programs, including Fashion Descience and Syndemic.

Fashion Descience is a runway show that connects researchers and fashion designers to create looks inspired by science. The organization is focused on "fostering collaborations between [the] two creative worlds and creating a new one that provides science with a new language and gives fashion

a new source of inspiration. It is about bringing together distinct perspectives, ideas, and developing a new mindset that has never been explored before."

In Fall 2014, Dr. Tal Danino (Post-Doc with Prof. Sangeeta Bhatia, MIT) and Tatiana Tejedor (Carmelo Clothing) participated in this runway competition as a team called 'Quorum 54'. Their design, shown to the right, used fabric printed with fluorescence microscopy images of bacteria-cancer cell interactions. In conversation with Dr. Danino, he explained that "the material was further enhanced by hand-painting part of the fabric with green fluorescent protein. This treatment makes the bacteria glow in the dark and helps visually illustrate their 'cancer-fighting' abilities.

The second project that Dr. Danino is working on is a microbiology video game that focuses on bacterial therapies against cancer. The game was initially developed by Alex May, a graduate student at VU University Amsterdam. He envisioned it as a tool to help "educate and empower people while explaining some of the mechanisms of cancer, like metastasis and tumor formation". The two researchers teamed up to create a real-to-life video game that has players (pyrogenic bacteria, chemotaxic bacteria, toxic-phage etc.) and power-ups (antibiotics, genetically engineered plasmids, etc.) to help the gamer defeat cancer/pathogens. This collaboration of Dr. Danino (U.S.) and Alex (Netherlands) has created something of an international platform for science outreach. In an interview with Alex and Dr. Danino, we were able to learn more about what Syndemic has to offer.

Dr. Danino and Tatiana Tejedor with model wearing 'Quorum 54' dress.
Photo Credit: Marta Murcia

MURJ: What was your inspiration for creating Syndemic?

Alex: "The main inspiration for Syndemic comes from the fact that despite the multitude of games that deal with destroying bad guys or blowing things up, I have never seen a game with pathogens/cancer as the actual enemy — just zombies and aliens, etc. If people enjoy fighting enemies, then I figure why not eliminate something that truly deserves it? Games typically have you fighting to save the world and our bodies are basically fighting pathogens all the time, saving our lives from countless threats. We are constantly in a war against disease: people need to be more aware of how powerful our immune system can be. Plus, it is empowering to be able to directly attack diseases with 'weapons', as opposed to letting some drug take care of it. My thinking is that by giving people a form of agency against disease, it can maybe make people feel a little more hopeful

about their own bodies and what they may be going through.”

MURJ: What is your vision for Syndemic in the future?

Alex: “Our ideal goal for the game would be to see it used as an introduction to educating people about pathogens, the immune system, and biotechnology. We would like for it to be a 'springboard' that hooks people into doing their own research/thinking about these sorts of topics. And while the game may not be entirely realistic, it can at least make people consider what is actually going on inside our bodies and hopefully allow them to have some fun doing it. I want people to ask, ‘Wow, is that what cancer actually does to our bodies?’ or ‘Is that what a virus looks like?’ or ‘How does the immune system actually protect us?’. We really like the idea of the game itself being critically examined by the players, and I plan to have messages that compare the real world to the game world.”

Dr. Danino: “In addition, we would like to collaborate with elementary and middle schools to use the game to help teach young children about basic biology. It could be a used as a visual tool to help illustrate the complex topics of disease, cancer, and immunology—similar to how The Oregon Trail video game is used for History classes. The video game, ideally, would generate a lasting enthusiasm for science.”

Dr. Danino’s multitude of science outreach projects—including Fashion Descience, Syndemic—all have one thing in common: they use visual media (clothes, videos, pictures, etc.) to demonstrate scientific topics. Dr. Danino firmly believes that “visual media stimulates learning andhelps people more readily believe and embrace new research ideas. In

addition, it can help show intuitive concepts without getting bogged down with too many details.”

Over the past few years, Dr. Danino has been part of several science outreach endeavors—one of his earlier projects is coolsciencevideos.com. The website hosts interesting science videos that focus on current research, biology topics, chemistry topics, activities for kids, etc. It serves as an easy way to learn advanced scientific ideas.

A currently developing project of Dr. Danino’s—in addition, to Fashion Descience, Syndemic, and coolsciencevideos.com—deals with creating porcelain plates printed with bacterial artwork. The project is in collaboration with artist Vik Muniz. The idea is to integrate art and science in a visually pleasing way. The plates integrate into everyday life and can bring positive appreciation to bacteria/microbiology.

To learn more about Dr. Tal Danino’s research please visit www.taldanino.com. ■

Examples of bacterial artwork printed on porcelain plates.
Credit: Tal Danino, Vik Muniz

MURJ Spotlight:
Professor David C. Page
Director of the Whitehead Institute

This issue's spotlight features Professor David C. Page, Director of the Whitehead Institute, Professor of Biology at MIT, and Investigator at the Howard Hughes Medical Institute

BY RIYA JAGETIA, ELIZABETH BERG, JAE HYUN KIM, BRIANNA JONES

Professor David C. Page, M.D., graduated from Swarthmore College with a degree in chemistry and received his M.D. from Harvard Medical School in 1984. Subsequently, Page was appointed a Whitehead Fellow and, two years later, was awarded a MacArthur "Genius" Grant. He joined the faculty of the Whitehead Institute and MIT in 1988 and is presently both a professor of biology at MIT and the director of the Whitehead Institute. In 1992, his lab mapped the human Y chromosome, and in 1997, Page was awarded an Amory Prize for advances in reproductive biology by the American Academy of Arts and Sciences. In 2003, his research group was able to sequence the human Y chromosome; the same year, Page also received the Curt Stern Award from the American Society of Human Genetics. In 2005, Page was elected to the National Academy of Sciences. In 2011, Page won the March of Dimes Prize in Developmental Biology. His lab at the Whitehead Institute currently focuses on the evolutionary and developmental origins of sexual dimorphism in the genome and germline to better understand the X and Y chromosomes as well as germ cells.

Professor David C. Page
All photos reproduced with permission from David C. Page archives

MURJ: You've had a long and illustrious career as a biologist. Going back to the beginning, what compelled you to study chemistry as an undergraduate?

Page: I had a lot of interests as an undergraduate. I arrived at college thinking that I was going to be an environmental lawyer. As a sophomore, I was briefly an economics major. Then I ended up as a chemistry major. I don't know

exactly why. I remained interested in all of these things, but I decided that this world of biological revolution was unfolding and was pretty interesting. I was actually a chemistry major and a biology minor, so I was very much in the biological space. For me, it was more of a decision about whether I would be in the natural sciences or if I would study something like economics or political science. I remain interested in all of these things. I briefly considered being a

religion major, just to round things out.

I would say, the older I get, the more I realize that I'm interested in understanding how the world works, at many different levels. This is partly why I'm drawn to studying the genetic differences between males and females. It's not hard to connect that to human life. Anybody can identify with these sorts of questions, and so I'm just interested in how it all works. If we can deal with particular little parts of these questions, that's very satisfying.

If I go back to when I was an undergrad in the mid-seventies, biology had not yet become molecular. It was just the beginning, the early days of that pursuit. Maybe today I'd be a biology major, but chemistry seemed to me at that

> *"It was the prehistoric days of the Human Genome project..."*

time to provide a more molecularly grounded approach to the kinds of things that I was thinking about.

MURJ: Studying chemistry in college was only the beginning of a very distinguished career. Can you give us a brief overview of your career?

Page: I had a bit of an odd history. I started doing research when I was an undergrad, and I decided to apply to both M.D. and M.D./Ph.D. programs. I got into a bunch of M.D./Ph.D. programs, and then someone told me that to complete an M.D./Ph.D. could

take as long as six years. I thought -- that's too long; I'm not willing to sign up for six years. So I came to the HST M.D. program. Four years -- fine. I had done a fair amount of research as an undergrad. In the HST program, you have to do some research to get the M.D. I started working in a lab the summer after my first year of medical school. Looking back upon it now, I realize that, in the summer of 1979, I became the first student anywhere in the world to work on what would become the Human Genome Project. I never heard the phrase Human Genome Project until seven years later, but I was working on it in 1979. I was there, before the beginning. It was just an incredible opportunity when I look back upon it. Of course, at the time, it didn't look like anything. I spent that first summer working in a lab at UMass Medical School, but I was working together with a professor who was then at MIT, David Botstein. It was the prehistoric days of the Human Genome Project.

I then went back to medical school doing clinical rotations, more or less thinking I'd become a cardiologist. But after completing my third year of medical school, I came back to finish up my HST thesis. That's when I decided to take a year's leave of absence. I planned to spend six months in lab and six months working in a hospital in a third-world country. One thing led to another. It took me longer to set up the third-world project than planned, and a one-year leave of absence became a two-year leave of absence.

I ended up spending three months working in a hospital in Liberia, where among other things, I met my future wife, who was a Canadian medical student working at the same hospital. So, that was a very important trip to West Africa! My medical education ended up being six years long by the time I was done. I paid my own way through medical school, and I didn't get a Ph.D. But, I spent more than two years, almost three years, working on what would become the Human Genome Project. I spent three months at a hospital in Liberia, where I met my future wife.

When I graduated from medical school, I was the only member of my graduating class who did not do an internship. Instead, six weeks after I graduated, the Whitehead Institute opened, and I became the first fellow here. Now, realize -- I didn't have a Ph.D. I hadn't done a post-doc, but somehow I was given this opportunity to run my own laboratory. Thirty-one years later, I'm still here! So that's the short version of my career.

I had a lot of decisions to make about what to do with my life. I did go to medical school, I didn't do a Ph.D., I didn't do an internship or residency. I couldn't possibly replay that path if I had to. I was sort of making things up as I went. I feel incredibly fortunate to have had a series of mentors and opportunities along the way. I always tell people to keep their eyes open and be prepared for that which they had not planned. Do things you're interested in, and keep your eyes open. You might meet your wife in Africa; you might stumble into the Human Genome Project.

MURJ: One of the greatest developments of our time has been the sequencing of the human genome, which you played an instrumental role in. With this development, medical care now has the potential to become greatly personalized. What are your thoughts on the ethical issues that may follow?

Page: Actually, I was educated in the "ethics" sphere more than 20 years ago. We were in the early stages of preparing to sequence the human genome, so we set up the Whitehead Institute Task Force on Genetics and Public Policy, bringing all sorts of people to campus to debate issues of genetic privacy, stem cell research, etc. What I've witnessed since then has been a tremendous generational transformation. I'm of a generation that is fairly obsessed with privacy. Your generation – as best as I can tell – has no idea of what privacy is.

MURJ: As evident through Facebook, Twitter, and the like.

Page: Right, right – as you might expect, I don't have a Facebook account. But, I think the shift is great. My generation, and even my parents' generation, was perhaps too focused on keeping

"I always tell people to keep their eyes open and be prepared for that which they had not planned...You might meet your wife in Africa; you might stumble into the Human Genome Project."

information to ourselves. My sense is that those days are gone.

In light of this transformation, it's interesting how none of the ethical issues have really worked out the way we'd expected them to. Back in the early 90's, we were worried that with the sequencing of the human genome, insurance companies would start discriminating against people based on their genotypes.

MURJ: As in the movie Gattaca?

Page: Yes! Contrary to our worries, such discrimination hasn't really happened. We realize that genetic determinism has its limits, as in we're not just the results of our DNA sequences, but also of our environments, lifestyles, etc. Obviously, there are cases like cystic fibrosis or sickle cell anemia where knowledge of the underlying DNA sequence provides strong predictive information, but for most of us, it's not nearly so simple. As we heard President Obama mention in his most recent State of the Union address, the era of Precision Medicine may address these individual complexities.

Going back to the question: Unlike 20 years ago, I'm not as concerned about privacy issues. Through my own experience, I've come to understand that public

policy concerns tend to morph and change over time. I think the questions now revolve more around: How do we feel about altering our genomes? Which reminds me of a New York Times article that I read last week on the call for a moratorium on CRISPR-Cas 9 editing of human germ line genomes. The first person quoted in the article, David Baltimore, was actually the founding director of the Whitehead Institute and the person who hired me in 1984.

Moving forward, I think we're going to have an endless stream of new ethical issues to consider. These issues will reinforce the need for all of us to continue becoming more broadly, as opposed to narrowly, educated. For we are not dealing with strictly scientific

questions, but rather, questions that require us to fall back on our understanding of how the world works and how it should work.

MURJ: A lot of the discussion ties back to knowledge, specifically of educating people about scientific, ethical considerations. What is your take on science education and/or communication for the general public?

Page: I believe it's the right, privilege, and responsibility of people who have been so lucky as to receive an MIT education to themselves become educators of everyone with whom they come in contact. It may be difficult for students to appreciate just how privileged they are – to be surrounded by incredible classmates, taught by fantastic professors, all in an environment where exciting things are happening a mere 100 feet away. And to realize that they're just a tiny, tiny fraction of the American population, of the world's population, that gets to experience

> *"Moving forward, I think we're going to have an endless stream of new ethical issues to consider....questions that require us to fall back on our understanding of how the world works..."*

first-hand what the debates are all about. So I'd like to take your question and turn it into a call to action on our part to share the knowledge we've been so fortunate to receive.

Because in my mind, there's no bound to where education starts or stops. There should be constant communication of knowledge. At the time when I started training as a scientist, many of my colleagues thought we had no particular responsibility to concern ourselves with ethical issues or public policy debates. Today, I doubt anybody feels that way. I think scientists now realize that it's our right and responsibility to engage with these sorts of issues – that, in fact, the future of the world depends upon our conscientious engagement.

Back in 2013, during my very first 7.016 lecture, I tried to take the students back to Cambridge in 1976, when the Cambridge City Council was holding hearings on whether to impose a 3-year moratorium on recombinant DNA research. I love to point this out to students because almost 40 years later they look around and see 150 tech companies based on recombinant DNA within a mile of our classroom. So who won that debate?

These issues are always going to be there. They aren't trivial issues, but it is our right, privilege and responsibility to help guide and educate people in whatever way we can.

MURJ: Going back to the theme of education, what do you like most about being a professor at MIT?

Page: It's very simple: The fantastic students. If you're a teacher, your happiness is determined by the quality of the students, and the students at MIT are great. They make being a professor a great pleasure, so I feel spoiled. I'm surrounded by great students at all levels – undergraduates, grad students – and then I have great faculty colleagues. MIT students are very smart, tend to be very idealistic, and very hard working. It's that combination of idealism, sincerity, earnestness and intelligence. It's perfect! I love that combination.

There is something also about the feel of this campus, of this community that is very focused on excellence, but also very down to earth. The reality is the students at MIT don't have any idea how good they are, and that's a wonderful characteristic.

MURJ: Many undergrads look up to MIT professors as their role models, for a variety of reasons. Who has been your greatest role model?

Page: I've had a lot of role models, but my strongest role models have to be my parents. They had great aspirations for me. They did not have the educational opportunities that I had. My mother completed 8th grade, and my father finished high school – but they taught me all that I needed to know about how to be a person.

I also had other role models as well. In college, I did some research at the NIH. I couldn't find a place to live, so the man I was working for, Bob Simpson, offered me his family's basement. So I actually lived with Bob's family, and over the course of a year and a half I was with them for a total of 12 months. Bob was an M.D.-Ph.D., and he went on into basic science, and he had a family, so he was a walking, breathing illustration of many things for me. He was also a

wonderful person. He was definitely an important role model for me.

Another one of my role models was David Botstein, who I trained with during medical school. He was fantastically intelligent, fantastically creative. He had high expectations of the people who were training in his lab but he was very hands-off, so I had a lot of freedom. He assembled this incredible group of students in his laboratory. There are so many people who were in that lab who are now distinguished professors all over the place. It was a great environment, a great place to train.

MURJ: Looking forward, what do you think the future of biology will be?

Page: Biology is the science of our time, and there is no end in sight to that. It's dangerous to make predictions about the future, but thinking about how it has changed over my professional lifetime, it has grown enormously. It was not the science of our time when I was in high school or college or even medical school. It has become that in the 30 years since. I think the sky is the limit.

"Biology is the science of our time, and there is no end in sight to that...I think the sky is the limit."

I do believe that even though biology has found so many ways to be applied, it's going to be absolutely critical that we maintain the places, and capabilities, and resources and such for people who want to do wildly creative things in understanding how living systems work. I think we're going to discover that. I think that we're going to have a lot more transformative discoveries that we can't even imagine. These are happening with increasing frequency, the kinds of discoveries that completely disrupt our understanding of how things work. There are many levels at which we have not come to any understanding of how they work.

MURJ: What do you think are some of the biggest current problems that biologists face?

Page: I'll tell you about one of the ideas I'm obsessed with, which is that 20 or 30 years from now, I imagine that we're going to understand at a very fundamental level the biological differences between human males and females. When we look back on today, where we have a kind of unisex human biology, we will see it as an odd, quaint, uninformed time. We'll look back on the present as a time of primitive understanding. ∎

MURJ UROP Summaries

Simulation of Blood Microcirculation and Study of Sickle Cell Disease

Jean-Pierre Theron[1], Raymond Jansen[1], Wesley Harris[2], Francois Le Floch[3]

[1]Student contributor, Class of 2015, Department of Aeronautical and Astronautical Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA

[2]Department of Aeronautical and Astronautical Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA

[3]MIT Alumnus of the Department of Aeronautical and Astronautical Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA

Abstract

Sickle cell disease is a genetic disease that affects the red blood cell behavior of a great number of people. In an effort to help the development of a cure, Dr Francios Le Floch previously developed a sophisticated mathematical model that simulates the flow and oxygen diffusion characteristics of capillaries. The aim of this project is to replicate the results that Le Floch acquired as well as to expand the calculated model by calculating intermediate data points. Through the use of Le Floch's model a preliminary set of three data points were calculated, of which the blood flow properties were consistent with those of Le Floch. The oxygen properties, however, did not converge by the end of the project nor were all the data points calculated. The research project will therefore be continued.

Background

Sickle cell disease is a form of Anemia where, upon deoxygenating, the red blood cells (RBC) deform into crescent-shaped cells. This is in part due because of the polymerization of the hemoglobin and the subsequent formation of long stiff fibers in the cytoplasm (Fiber formation, 2004). These sickle cells then affect the flow properties of the blood flow by, for example, blocking capillary vessels, which can then lead to severe pain and damage of organs.

In an attempt to help research for a cure, Dr Francois Le Floch, as part of his doctoral research and in conjunction with Prof Wesley Harris of the Department of Aeronautics and Astronautics, created a mathematical model that simulates the flow characteristics of sickle cell blood flow in capillary vessels. Through the use of numerical methods like the Immersed boundary method, the model calculates a large set of flow and oxygen diffusion equations (which include the likes of the Navier Stokes equations) to get the necessary results in terms of oxygen transfer, blood plasma dynamics, the mechanics of red blood cells and many more physiological constraints (Le Floch, 2010, pp. 75-76). The hope is that researchers could use this model to test treatment options since they can now get an estimate of how the flow properties are affected by their proposed treatment (Le Floch, 2010, pp. 160-176).

The aim of this research project was thus to model different case studies of different flow conditions. This included replicating the results of Dr Le Floch's PhD thesis and modelling completely new scenarios.

Method

The case studies that were analyzed were grouped together under different overall scenarios. These scenarios were distinguished from each other based on their specific flow parameters, for example different hematocrits or fixed inlet blood flow velocity. The case studies themselves were then distinguished based on only one parameter, capillary diameter. The range of capillaries that were analyzed started at a diameter of 10 μm and goes down 5 μm. In Le Floch's research, data points were analyzed at intervals of 0.5 μm. In the

research project intermediate data points that fall half way between Le Floch's data points (Le Floch, 2010, p. 103) were also analyzed. Hence intervals of 0.25 μm were analyzed during the research project.

The approach that was taken consisted of feeding predetermined flow conditions (gained from research done by Le Floch) as input to the model. The model would then run until the calculations converge. Convergence was achieved when the red blood cell geometry hasn't changed by more than 1%, with the two points of comparison being at least twice the spatial period apart (Le Floch, 2010, p. 63). The spatial period was indicative of the distance traveled by the RBC in the capillary and was defined as:

$$L = \frac{V_{RBC}}{\pi(Hct)(R)^2}$$

where L is the spatial period, V_{RBC} is the RBC volume, Hct is the Hematocrit and R is the capillary radius. After a set of calculations had finished for a particular vessel diameter, the resulting red blood cell geometry was used as input to the subsequent case study with a smaller capillary diameter. This process was then repeated for the entire range of capillary diameters.

Results

By the end of the research project, the calculations for all the case studies across the various scenarios had not been completed. Preliminary results, however, did look promising as the blood flow property results of three case studies (10 μm, 9.5 μm and 9.0 μm capillary diameters) in a normal blood flow scenario, where inlet blood flow velocity was fixed, were replicated. The calculations converged within the 1% margin that was set before hand and the error between the newly acquired data and the data that was acquired during Le Floch's research was lower than 5%. The new intermediate data point, which was calculated, also fitted within the previously established trends, which validates its values.

Figures 1 to 3 provide a visual representation of how the original data and calculated data compare. Figure 1 shows how the replicated/calculated discharge hematocrit (the ratio of the volume occupied by the RBCs to the total volume of the blood flow) compares to the original data calculated by Le Floch. Figure 2 plots RBC velocity (the velocity at which the RBC travel in the blood stream) over the capillary sizes (graphs A). Figure 3 plots apparent viscosity

Figure 1. Discharge haematocrit at steady state vs. capillary vessel radius for the original data of Le Floch and the replicated/extended data.

Figure 2. RBC velocity at steady state vs. capillary vessel radius for the original data of Le Floch and the replicated/extended data.

Figure 3. Apparent viscosity at steady state vs. capillary vessel radius for the original data of Le Floch and the replicated/extended data.

25

(resistance to shear deformation of a non-Newtonian fluid) of the various capillary radii

Additionally a graph of available data (Le Floch, 2010, pp. 111-112) of a scenario that studied blood flow with a lower oxygen partial pressure (PO2 = 40mmHg) was plotted on the three figures so as to give an idea of how the scenario that was studied compared to other similar scenarios. On Figures 2 and 3 additional available data (op. cit.) of another scenario, which used a constant inlet pressure boundary condition, were plotted (graphs B) to show how a change in boundary conditions affects the data.

Despite the convergence of the blood flow properties, the blood oxygen properties of the three case studies had not converged and therefore these calculations will have to be continued.

Future Work

As stated before, there are still a lot of case studies that need to be run. Furthermore the calculations for the blood oxygen properties of the previously mentioned case studies will also be continued to achieve convergence. The hope is also that the additional data points that are generated will greatly help expand the model to make it even more representative of the blood flow dynamics of sickle cell patients. Ultimately, however, the expectation is that this expanded model will help medical researchers in their goal of developing a drug or technique to help those who are afflicted with sickle cell disease.

References

Fiber formation. (2004, March 05). Retrieved November 03, 2014, from About Sickle Cell disease: http://www.sicklecellinfo.net/fiber_formation.htm

Le Floch, F. T. (2010). Design of a Numerical Model for Simulation of Blood Microcirculation and Study of Sickle Cell Disease (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, United States of America.

MURJ
Reports

Gathering Real World Data to Predict Admission Statistics

Noor Khouri[1], Dimitrios Pagonakis[2], Marta González[3]

[1]Student Contributor, Class of 2015, Department of Civil Engineering, MIT, Cambridge, MA 02139, USA

[2]Student Contributor, Class of 2015, Department of Civil Engineering, MIT, Cambridge, MA 02139, USA

[3]Human Mobility and Network Lab, Department of Civil Engineering, MIT, Cambridge, MA 02139, USA

In this paper we design a model that estimates the incoming flow of international graduate and undergraduate students at the Massachusetts Institute of Technology (MIT). The authors, both international students at the university, were motivated to explore this question due to their diverse cultural backgrounds. The analysis uses a modified version of the Gravity Model used in network theory and estimates the number of students coming from each country, given only the country's data, namely the distance from MIT, the Gross Domestic Product (PPP) per capita and the population size. The model parameters were optimized using student data from the MIT International Students Office (ISO) for the 2013-2014 academic year, but could be further refined by using data from previous and later years. This model could be proposed to any International Students Office in the US as a means of diversifying the respective university's student communities and to provide an aid for setting country specific admission quotas.

1 Introduction

The MIT student body consists of 11,300 undergraduate and graduate students. Among them, approximately 3,100 are international students: 400 undergraduates and 2,700 graduate students. For the 2013-14 academic year, these students come from 110 different countries of various cultures and backgrounds. Given that there are 192 countries, it may appear as though the MIT student body is a relatively broad representation of the world, but is this the case? Are countries with financial or political problems underrepresented in the MIT community?

The analysis presented in this paper addresses these questions by proposing a Gravity Model approach to predict the total expected number of MIT students from a given country based on parameters such as the distance of a country to MIT and economic measures, including population size and GDP per capita. Using networks as a visualization and data analysis tool, we construct two parallel networks: the World Network and the MIT International Students Network, wherein each node represents a country that is linked to its bordering countries. The World Network consists of 192 nodes (countries) and 300 edges (political borders), whereas the MIT Network consists of 110 nodes and 146 edges, a subset of the World Network.

The following section describes the construction of the networks, including the details of each node attribute, in conjunction with the development of the model as a predictive tool.

1.1 Bordering Networks

Borders across countries tend to provide a natural burden for community network expansions (Cerina, Chessa, Pammolli, & Riccaboni, 2014). In other words, the formation of network communities is significantly influenced by country borders and the number of neighboring countries; this property also manifests itself in the tendency for international students at MIT to form clusters with others from nearby nations. As such, the empirical networks we construct are visualized by connecting each country (nodes) with its nearest neighbors.

The motivation to connect the nodes with links representing political borders is further justified by the analysis conducted by Cerina et al (2014). Figure 1 plots the normalized distance distribution of the links between countries in Europe and between states in the US. The observed rapid decay in the European probability distribution suggests that the spatially embedded networks that form in the US span multiple states across the whole country without any characteristic geographic

scale, while the analogous European community connections correspond to neighboring nations and contiguous regions (Cerina et al., 2014).

Figure 1. Normalized distance distribution for Europe (blue dots) and the US (red dots) and their best fits (Cerina et al., 2014). The connectivity within the US follows power law decay with respect to distance while Europe exhibits an exponential decay.

Furthermore, Figure 1 demonstrates that European countries' community networks are transnational and tend to exhibit large clustering; communities are very close to one another and are at most linked with their direct bordering countries. In other words, the probability of an Italian knowing someone from Denmark is a lot lower than for the case of a Californian resident knowing a person from Massachusetts. This illustrates the important insight that communities in the US are well spread and well connected as compared to the tightly packed, but less connected, European communities that are dependent upon geographic proximity and contiguity. Furthermore, although the gravity model analysis is independent of the nodal connectivity, the physical borders play an important role in the network visualization and the resulting inferences regarding the diversity of the MIT student population.

1.2 The Gravity Model

To build a model that uses country data as inputs and returns a prediction of the number of students at MIT from each given country, we modify the traditional gravity function that is often used in transportation applications to forecast travel demand. This gravity function, which is analogous to Newton's theory of gravity, measures the attraction between two masses in terms of their population sizes and their separating distance. More specifically, the generic gravity definition is given below.

Definition 1 *Let locations i, j have populations P_i and P_j, respectively, and let C(i,j) represent the travel cost friction factor between them. Note that C(i,j) depends on the distance $r_{i \to j}$ between i and j. We define*

$$T_{i \to j} = \alpha \frac{P_i{}^b P_j{}^c}{C(i,j)^d}$$

to be the gravity from origin i to destination j, with fitting parameters a,b,c,d∈R .

To estimate the size of a given country's community at MIT, it is clear that its financial state would affect the chance of its students being able to study at MIT. Hence, for our model we redefine the travel cost friction factor to account for both the distance and GDP (PPP) per capita of each country relative to the US. Thus, our model of gravity will predict the population of students at MIT from country of origin i by the formula

$$T_{i \to MIT} = \delta \frac{P_i{}^\alpha P_{USA}}{r_{i \to MIT}{}^\beta \left(\frac{GDP_{USA}}{GDP_i}\right)^\gamma}$$

P_i is the population of country i, $r_{i \to MIT}$ is the distance from country i to MIT, and GDPi is the Gross Domestic Product (PPP) per capita of country i. The parameters α, β, γ, δ are fitting parameters which shall be determined by optimizing the model using the empirical data for the number of students from each international country that is represented at MIT.

Note that this gravity model assumes that the sample of the population from which students can be selected for admission to MIT have fulfilled the academic requirements and have attained the necessary qualifications. Prior to optimizing the model with the actual MIT student data, we present the sources and organization of data used to perform this analysis.

1.2 Data

In order to construct the MIT and World networks and to develop the gravity model for predicting the number of international students from a particular country who study at MIT, we began by collecting the necessary data from a variety of sources.

The international students' data for the 2013-2014 academic year, categorized by country, was obtained by contacting the International Students Office at MIT

(Cheng, 2014). The country population data analyzed in this paper was extracted from The World Bank ("Data - Population, total," 2013). The GDP (PPP) per capita of each country was derived from the CIA World Factbook ("The World Factbook - GDP Per Capita (PPP)," n.d.), while the distance of each country to MIT was estimated implementing the Haversine formula (Sinnott, 1984). This formula outputs distance in terms of the longitude and latitude of two points assuming that the Earth is a perfect sphere (which is a much better approximation than Euclidean formula).

Thus, each country in both networks was assigned with the following unique node attributes:
- Country name (Label)
- Number of undergraduates at MIT
- Number of graduates at MIT
- Total number of students at MIT
- Total population
- GDP (PPP) per capita
- Average longitude
- Distance from MIT (using the Haversine formula)
- Predicted number of students at MIT (based on results of gravity model)

As previously mentioned, the connectivity of the networks was defined based on countries' borders. The bordering country data was obtained from the CIA World Factbook and the links were created via an adjacency matrix (Barabási, 2012, p. 31; "The World Factbook - Land Boundaries," n.d.). In the next section we present the methods of analysis of the gravity model.

2 Methods and Results

Upon optimizing the gravity model using the existing international MIT students' data, we were able to determine the fitting parameters necessary to predict the number of incoming students from a particular country for the upcoming academic year. This section evaluates the strength of the correlation between the distance of a country to MIT, the country's population and its GDP per capita and the number of students within the MIT international student body who originate form the country in question. Finally, conclusions are drawn regarding the relative impacts of each parameter on the number of students attracted to MIT from a given country.

2.1 Distance of a Country to MIT

Following the generic gravity model, we first explored the relationship between the distance of a country from

MIT and the number of students from that country who are currently studying at MIT. This was primarily done to gauge whether or not the distance of the country in question would be a good indicator for the number of students it could send abroad to pursue their higher education at MIT. Figure 2 shows the results of the model predictions for the number of students that could be admitted to study at MIT plotted versus the distance from each country to MIT.

Figure 2. The box plots show the predicted number of students at MIT for given distances from MIT. Note that the upper and lower bounds of the boxplots show the quartiles, the minimum and maximum predicted values, while the red line represents the median and the whiskers extend to the extreme outlier data points. The raw data is shown as individual points (green), each representing a particular country.

The predictions were derived from the node attributes of the countries comprising the MIT International Students Network for the academic year 2013-2014. The raw data from the MIT Network is shown as individual data points. As made evident in Figure 2, there is no clear correlation between the distance of a country to MIT and the number of students at MIT of that origin. Both data sets fluctuate significantly and do not exhibit any particular trend. In fact, each box plot, representing the model predictions for a logarithmic range of distances, covers a variety of values for the predicted number of students at MIT. This large variability could be explained by the ease of modern air transportation, which facilitates traveling thousands of miles to an accessible location, such as Boston. As a result, the distance of the country of origin to MIT is the least important factor that contributes to the inflow of new students; students who

are geographically closer to MIT are, in fact, not more likely to study at MIT than others who live further away, as illustrated by Figure 4. Indeed, this notion is justified by the optimized fitting factor β, which was estimated to equal $2.5\text{-}10^{-2}$, the smallest of the fitting parameters.

2.2 Population of a Country

Next, we examined the relationship between the population of a country and the number of its students at MIT. Figure 3, which highlights the countries with the largest populations in the world, and Figure 4, which shows the distribution of the international students at MIT, can be compared qualitatively. Although the country populations do not perfectly correlate with the size of MIT student populations, there is clearly a relationship between the two parameters.

In particular, both China and India dominate in both the MIT and World Networks. Following on from the qualitative analysis, we then plotted the model-derived predicted number of students at MIT versus the population size of each country. The model results and the raw data (individual cyan data points), shown in Figure 5, are closely matched and exhibit an increasing trend. Figure 5 clearly shows that there is a direct correlation between the population size of a country and the number of students originating from the country in question. This can be rationalized by the fact that the larger a country's population is, the more likely it is to send students to MIT given the larger sample size of potential qualified applicants, and thus the expected number of students that can be admitted increases. It is worth noting, however, that according to the model we developed, certain countries are either over-represented (raw data points lie above model predictions) or under-represented (raw data points lie below model predictions). For example China, which has the largest populations in the MIT and World Networks, still has an over-represented population at MIT according to our model. Furthermore, we found the fitting parameter of the population α, to be $6.4*10^{-1}$, the largest of the fitting parameters in the gravity model. Hence it represents a best estimate of the number of students at MIT compared to the other input parameters considered in the proposed model.

Figure 3. This visualization represents the empirically constructed World Network wherein the countries (nodes) are ranked by both color and size based on the sizes of their respective populations. The red nodes, representing China and India, have the largest populations relative to the other countries in the world. Green nodes, such as the US and Brazil represent the next tier of large population sizes, while the gray nodes have comparatively smaller populations

Figure 4. This visualization represents the empirically constructed MIT Network wherein the countries are ranked based on the sizes of their respective populations. China, shown in red, has the largest international student population at MIT, while India, Canada and South Korea, shown in orange, also exhibit significant representations.

Figure 5. The box plots show the gravity model predicted number of students at MIT for given country population sizes. The model predictions closely match the raw data and exhibit positive correlation with the number of international students from a particular community at MIT. The population of a country appears to be the best measure of the MIT student population from the parameters considered in the proposed gravity model.

2.3 GDP (PPP) Per Capita

Given that Oct4 expression is a key property of naïve plFinally, in order to account for the financial situation of a country and to determine its effect on students' abilities to study at MIT, we examined the distribution of GDP per capita across all countries in the world; the results are shown in Figure 6. According to the visualization, if GDP were to significantly impact a student's ability to study at MIT, countries in Europe and certain regions in Asia would be more representative within the MIT international student body.

To test the effect of finance on the size of a country's student population at MIT, we proceeded with a quantitative analysis in the same manner as in the previous sections. The number of students at MIT was plotted versus the GDP per capita for both the raw data and model predictions. Comparing the predicted population of our model with the actual data, we see that there is a close correlation between a country's financial status and the number of its students that receive higher education at MIT. This naturally makes sense since private institutions like MIT are not easily accessible to developing nations. The outliers of the model located above the predicted data represent countries of low GDP that potentially sent students to MIT thanks to the availability of financial aid. Note though that even these countries have an acceptable GDP, which presumably means that its students are educated enough to be aware and capable of going through the university and financial aid application process. However, students from Least Developed Countries rarely have this opportunity. The outliers located below the model data tend to lie on the high side of GDP and represent countries where students do not have any incentive to study at a US institution given that they have access to established universities in their countries (i.e. European countries, Singapore, Japan). Additionally, pursuing higher education in their countries could be less of a financial burden due to reduced tuition fees for nationals, as in UK. Finally, this could perhaps be justified further for countries such as Qatar, which actually has the highest GDP in the world that might not have a culture of sending students to study abroad. In any case, the relationship between GDP and the country's population at MIT is still notable, as evidenced by its fitting parameter γ, which was $6.0*10^{-1}$.

Figure 6. This visualization represents the empirically constructed World Network wherein the countries are ranked based on their GDP (PPP) per capita. Larger, red nodes represent countries with a high GDP, whereas smaller, gray nodes represent countries with a lower GDP. Countries with high GDP tend to cluster in North America, Europe, the Arabian Gulf and within certain portions of the Asian continent.

Figure 7. The box plots show the gravity model predicted number of students at MIT for given country GDP (PPP) per capita values. The model predictions quite closely match the raw data and exhibit positive correlation with the number of international students from a particular community at MIT. The GDP per capita of a country appears to be a good measure of the MIT student population.

Figure 8. This visualization represents the predicted MIT student populations, wherein the countries are ranked based on their predicted representations at MIT. The gravity model predicts higher representations from European countries as well as Brazil and Japan, as compared to the existing MIT populations shown in Figure 4.

2.4 Gravity Model Results

The Gravity model was quite successful in accurately predicting the approximate number of students at MIT from each country in the world given its demographic data. The gravity formula can be further enhanced by adding additional parameters and existing data to improve its accuracy; suggestions for such variables are discussed in the concluding remarks. Given only 110 existing data points from the MIT Network, the model's estimates are quite impressive. Figure 8 shows the number of students from each country, as predicted by the model, while Figure 4 illustrates the actual data for the 2013-2014 academic year; the great contributors, such as China, India and Canada are common between the two visualizations, however, the model predicts larger student representation from European countries. In addition, there are several outliers which are either over or under-represented, for reasons the proposed gravity function does not capture, namely political and societal influences, that are perhaps more difficult to quantify. Table 1 provides a comparison

between the raw data and model predictions for sample countries that illustrate the model's accuracy. Countries, such as China are over-represented according to our model, while others, such as Iraq, are under-represented, most likely due to political instabilities with the US. Furthermore, Table 2 shows the variance of each of the fitting parameters that constitute the proposed gravity model. Evidently, population, with the smallest fitting parameter variance, has the most significant contribution in predicting the number of students within the MIT international student community.

3 Conclusion

Through a series of qualitative analyses, which were facilitated by the various network visualizations, in addition to the quantitative gravity model formulation, the composition of the MIT student body has become clearer. Indeed, as shown by Figures 4 and 5, international students at MIT represent 110 countries of the possible 192 countries comprising the World Network and thus form a

Country	Actual Students	Predicted Students
Argentina	19	15
Bahrain	2	2.17
Ecuador	5	5.76
France	87	30.52
Greece	46	7.49
Haiti	1	1.24
India	314	52.27
Iraq	1	7.17
Turkey	56	19.47
UAE	4	7.6
UK	56	30.78
Venezuela	15	10.31

Table 1. Note that the countries in green are under-represented at MIT due to factors that are not accounted in the gravity model (political, societal), while the countries in red are countries which are over-represented at MIT. Possible reasons for this could include the fact that alumni or close relations between the Institute and the country in question attract more students than anticipated by the model.

Population	GDP	Distance
$3.3 \cdot 10^{-2}$	$7.2 \cdot 10^{-3}$	$5.2 \cdot 10^{-2}$

Table 2. The variance table for the fitting parameters for Population (α), GDP (γ) and Distance (β). The lower the variance the tighter the spread of the data and hence the greater the influence of the variable to the model.

decently culturally diverse community. Nevertheless, the international student population at MIT is not an accurate representation of the World Network. It is also important to note that the most noticeable under-representation in the MIT network are countries in the Western African continent, which are also predicted to have a very low student population at MIT according to the results of our model shown in Figure 9. Furthermore, we observed that both the population of a country and its GDP are good predictors for the number of students at MIT, while countries with existing MIT students and alumni have a greater tendency of being attracted to the university. Perhaps this could be explained by MIT's reputation among the country's community or due to donations by alumni to the institute.

As our proposed gravity model for predicting the number of students that would be admitted to MIT

provides the potential for promising results, such study could have many applications. For example the collaboration between the International Students Office at MIT and the admissions office in order to improve the diversity of its incoming international student body. The statistics by the model could also be used to set admission quotas for various countries in the world.

In order to further refine the gravity model we have used to predict the number of students from a given country being admitted to MIT, a number of additional factors would need to be introduced into the model. Firstly, it would be important to build the model based on ISO data from several years and hence improve the accuracy of the model. Also, since the international students data provided by the MIT ISO is categorized according to the students' nationalities, we recommend that additional student data be incorporated into the

Figure 9. This visualization represents the predicted MIT student populations based on the World Network, wherein the countries are ranked based on their predicted representations at MIT. The World Network predictions closely match the MIT Network predictions shown in Figure 8, suggesting that MIT is more likely to attract countries that already have an established student population at the university. Countries that are missing in the MIT Network are predicted to send few to none of its students to MIT.

node attributes. This is due to the fact that a student's nationality does not necessarily reflect the conditions of the country from which the student applied. Thus, it would also be interesting to explore the effect of additional economic measures of development that shed light on the strength of a country's educational system, its political stability and its relations with the US. Finally, since this model accounts for the total MIT student population from a given country, dividing the MIT network into graduates or undergraduates could perhaps highlight other trends that are not captured in this simplistic version of the gravity model. For example, graduate students could potentially exhibit larger concentrations of particular populations from which professors at MIT have maintained collaborative relationships with alumni originating from the country in question.

Overall, we believe that the results presented in this paper highlight some interesting aspects of the MIT international students community and allow for greater avenues for future research.

References

Barabási, A.-L. (2012). Network Science. Retrieved from http://barabasilab.neu.edu/networksciencebook/

Cerina, F., Chessa, A., Pammolli, F., & Riccaboni, M. (2014). Network communities within and across borders. Scientific Reports, 4. http://doi.org/10.1038/srep04546

Cheng, E. (2014). MIT International Student Body Data. MIT Interntional Students Office.

Data - Population, total. (2013). Retrieved from http://data.worldbank.org/indicator/SP.POP.TOTL?order=wbapi_data_value_2013+wbapi_data_value+wbapi_data_value-last&sort=asc

Sinnott, R. W. (1984). Virtues of the Haversine. Sky and Telescope, 68(2), 159.

The World Factbook - GDP Per Capita (PPP). (n.d.). Retrieved from https://www.cia.gov/library/publications/the-world-factbook/rankorder/2004rank.html

The World Factbook - Land Boundaries. (n.d.). Retrieved from https://www.cia.gov/Library/publications/the-world-factbook/fields/2096.html

Improving the Efficiency of Carbon Fixation in *Ralstonia eutropha* with Carbon-concentrating Microcompartments

Sophia Li[1], Jingnan Lu[2], Christopher J. Brigham[3], Anthony J. Sinskey[4,5,6]

[1]Student Contributor, Class of 2015, Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

[2]Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

[3]Department of Bioengineering, University of Massachusetts Dartmouth, 285 Old Westport Road, North Dartmouth, MA 02747, USA

[4]Department of Biology, [5]Division of Health Sciences and Technology, [6]Engineering Systems Division, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

Microorganisms such as *Ralstonia eutropha* offer renewable solutions to current global warming and carbon emission challenges with their ability to use CO_2 as the carbon source for valuable products such as biodegradable plastic and biofuels. An important bottleneck in this process is *R. eutropha's* carbon fixation efficiency, which is limited by the enzyme, ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCO), since RuBisCO can take both CO_2 and O_2 as substrates. Cyanobacteria and some chemoautotrophs have increased the efficiency of RuBisCO by encapsulating it in carboxysomes, protein-based organelles that concentrate CO_2 while blocking O_2 from RuBisCO. Since these structures are not found in *R. eutropha*, we introduced carboxysomes into *R. eutropha* to improve carbon fixation and evaluated the carboxysomes' effect on bacterial growth and product formation. We found that the carbon-fixation enzymes found in carboxysomes, carbonic anhydrase and RuBisCO, had higher enzymatic activities in strains with carboxysomes as compared to a control strain without carboxysomes. We also determined that carboxysomes in the *R. eutropha* strain H16 led to greater cell growth, but decreased product formation of the polymer PHB when the bacteria were grown autotrophically on CO_2 or formate. Thus, this study shows the utility of carboxysomes in improving carbon fixation in nonnative microorganisms, while providing insight into potential challenges, such as carboxysome assembly, in increasing the yield of products that aggregate in the cytoplasm of *R. eutropha*.

Introduction

Climate change and the inevitable depletion of fossil fuels has led to a search for sustainable energy sources. Certain microbes, such as *Ralstonia eutropha*, offer a solution to these global warming and energy problems with their ability to grow rapidly, take in CO_2 as a carbon source, and produce a wide array of useful products (Lee et al., 1999; Lu et al., 2012).

R. eutropha is a Gram-negative bacterium that is able to use renewable carbon sources such as plant oils (Budde et al., 2011), organic acids (Yang et al., 2010),

CO_2 (Pohlmann et al., 2006), and formate (Li et al., 2012) to naturally produce polyhydroxybutyrate (PHB), a biodegradable polyester. In addition, *R. eutropha* has been engineered to produce other valuable products such as isobutanol and 3-methyl-1-butanol (Lu et al., 2012), which can be used as gasoline substitutes (Atsumi et al., 2008) or as chemical feedstocks (Gogerty et al., 2010). A *R. eutropha* mutant strain has been shown to produce over 14 g/L of isobutanol and 3-methyl-1-butanol after 50 days of flask cultivation (Lu et al., 2012). However, increases in yield and productivity are needed to make the production of these alcohols feasible on an industrial scale (Lee et al., 2008).

One of the bottlenecks in *R. eutropha's* ability to produce valuable products is its carbon fixation efficiency. While *R. eutropha* has several carbonic anhydrases (Figure 1), a key carbon fixation enzyme, ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCO), is inefficient because it takes both CO_2 and O_2 as substrates (Keys, 1986). Thus, in this study, we aim to improve the efficiency of carbon fixation in *R. eutropha* by introducing carboxysomes, a carbon-concentrating mechanism, in the bacterium.

A carboxysome is a bacterial microcompartment that is similar to a synthetic organelle (Yeates et al., 2008). It consists of an icosahedral protein shell surrounding the enzymes carbonic anhydrase and RuBisCO, separating them from the rest of the cytoplasm (Bonacci et al., 2012). Carboxysomes are natively found in cyanobacteria and some chemoautotrophic bacteria (Price et al., 1991), which often grow in environments with low concentrations of CO2. It has been hypothesized that the carboxysome increases the efficiency of carbon fixation through two main mechanisms: saturating RuBisCO with CO_2 and blocking RuBisCO from taking O_2 as a substrate for the production of wasteful byproducts (Figure 2; Yeates et al., 2008). By increasing the proximity of carbonic anhydrase and RuBisCO and trapping CO_2 inside the carboxysome, the specificity and efficiency of RuBisCO increases as it is saturated

Figure 1. Possible roles of carbonic anhydrases (red) in *R. eutropha*. Periplasmic Caa preferentially converts CO_2 into bicarbonate. Cytosolic Can supplies CO_2 to RuBisCO in the Calvin-Benson-Bassham (CBB) cycle. The roles of Cag and Can2 are less clear; Cag could play a similar role to Can while Can2 could be involved with pH homeostasis.

RuBisCO=ribulose-1,5-bisphosphate carboxylase oxygenase

Figure 2. Carbon-concentrating mechanism of the carboxysome. CO_2 and bicarbonate (HCO_3^-) enter the cell as described in Figure 1. Carbonic anhydrase (CA) and ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCO) are also inside the protein shell of the carboxysome. Bicarbonate enters the carboxysome through pores and is converted to CO_2. RuBisCO then catalyzes the carbon fixation reaction where ribulose-1,5-bisphosphate (not shown) and CO_2 are converted to two molecules of 3-phosphoglycerate (PGA). The carboxysome shell (purple) restricts oxygen from entering and interacting with RuBisCO to form the toxic side product, phosphoglycolate (PG). CBB cycle= Calvin-Benson-Bassham cycle.

with more CO_2. The protein shell also shields RuBisCO from O_2, minimizing the substrate competition between O_2 and CO_2 for RuBisCO, so less phosphoglycolate, a potentially toxic byproduct, is formed and the efficiency of carbon fixation is increased. Through these two mechanisms (Rae et al., 2013), the carboxysome allows the bacteria to more efficiently fix CO2.

Since carboxysomes have been seen to improve carbon fixation in many bacteria due to their carbon-concentrating mechanisms, in this study we introduced carboxysomes from the chemoautotrophic bacterium *Halothiobacillus neapolitanus* into *R. eutropha*. We then measured the carboxysomes' effect on the

enzymatic activity of the two key carbon fixation enzymes, carbonic anhydrase and RuBisCO. We also determined whether the introduction of carboxysomes improved *R. eutropha's* ability to grow on carbon dioxide/formate. Finally, we evaluated whether the addition of carboxysomes led to changes in the yield of the biodegradable plastic, polyhydroxybutyrate, in *R. eutropha.*

Methods

Chemicals, bacterial strains, and plasmids

Chemicals were obtained from Sigma Aldrich unless otherwise noted. Bacterial strains and plasmids used in this study are included in Tables 1 and 2.
Growth media and cultivation conditions

Table 1. Strains used in this work

Ralstonia eutropha strains	Genotype	Reference
H16	Wild-type gentamycin resistant (Gen^r)	ATCC17699
Re2061	H16 ΔphaCAB (Gen^r)	Lu et al., 2012

Table 1 continued. Strains used in this work

Escherichia coli strains	Genotype	Reference
One Shot Top 10 chemically competent cells	Strain suitable for high efficiency transformation	Life Technologies
S17-1	Conjugation strain for transfer of plasmids into *R. eutropha*	Simon et al., 1983

Table 2. Plasmids used in this work

Strain	Genotype	Reference
pBBad	L-arabinose inducible vector, mob, P$_{lac}$, lacZα, Kan^r P$_{BAD}$, T$_{rrnB}$, araC	Fukui et al., 2009
carboy	pBBad derivative with carboxysome operon	This work

R. eutropha strains were grown in either tryptic soy broth without dextrose (TSB) (Becton Dickinson, Sparks, MD) or minimal media (Lu et al., 2012) at 30°C with 10 μg mL^{-1} of the antibiotic gentamicin. *R. eutropha* strains with plasmids were grown with an additional 300 μg mL^{-1} kanamycin. Carbon sources used in minimal media cultivation were 2% (w/v) fructose or 0.2% (v/v) formic acid.

For enzymatic assays, a single *R. eutropha* colony from a TSB plate with 1.5% agar and 10 μg mL^{-1} gentamicin was used to inoculate 5 mL of TSB medium with appropriate antibiotics. After 48 hours of growth, these cultures were used to inoculate 50 mL flask cultures of minimal media with appropriate antibiotics. Flask cultures were continuously shaken at 200 rpm in a 30°C incubator. After one day of growth, the cultures were induced with 0.2% (w/v) of L-arabinose. Aliquots were removed at various time points for analysis.

When *R. eutropha* was grown with formic acid as the sole carbon source, 1.5 mL of the TSB medium culture was used to inoculate 150 mL of flask cultures. The pH and the OD$_{600}$ of the flask cultures was measured periodically. Whenever the pH of the flask cultures was measured to be 8, formic acid was added to the flask cultures (0.1% final concentration v/v) and the pH was adjusted back to 7. Flask cultures were induced with 0.5% (w/v) of L-arabinose during the exponential growth phase of the bacteria.

Escherichia coli strains were grown in LB medium (Bertani, 1951) at 37°C. *E. coli* strains with plasmid were grown with 50 μg mL^{-1} kanamycin.

Plasmid and Strain Construction

Standard molecular biology techniques as described previously were used for the construction of all plasmids and strains (Sambrook and Russell, 2001). Phusion DNA polymerase (New England Biolabs, Ipswich, MA), was used for amplification of DNA. The QIAquick gel extraction kit (Qiagen, Valencia, CA) was used to gel purify all DNA, while the QIAprep spin miniprep kit (Qiagen, Valencia, CA) was used to extract plasmids. All restriction enzymes used were supplied by New England Biolabs (Ipswich, MA).

To construct pCarboxy (Table 2), the carboxysome operon from *H. neapolitanus* was amplified through PCR. The forward primer (5'ATTACCCGGGTTGTACACGGCCGCATAATCG3') and the reverse primer (5'ACTACCCGGGCTAGCTAATGCTCTGTCTCA3') were designed with XmaI restriction sites. After amplification, the carboxysome operon was purified and digested with XmaI, forming a DNA product with a final length of 9811 bp. The pBBad vector (Fukui et al., 2009) was also digested with XmaI. The digested vector and carboxysome operon DNA insert were ligated together and transformed into *E. coli* Top 10 chemically competent cells (Life Technologies, Carlsbad, CA). Plasmids were then extracted and correctly constructed plasmids were found with a SalI diagnostic digest. The correctly constructed plasmids were then transformed into *E. coli* S17-1 (Simon et al., 1983), a donor for the conjugative transfer of mobilizable plasmids, and the plasmid was introduced into *R. eutropha* via conjugation using a standard mating procedure (Slater et al., 1998). *R. eutropha* with the plasmids were selected using

colony PCR with the same primers used to amplify the carboxysome operon.

Preparation of cell extracts for enzymatic assays

In order to prepare cell extracts of *R. eutropha* strains for enzymatic assays, 10 mL of cell culture were removed and spun down at 4000 x g. The cell pellets were resuspended in 2 mL of phosphate buffered saline. 1 mL of resuspended cells with added 0.1 mm zirconia/silica beads (BioSpec Products, Bartlesville, OK) were shaken five times at 5.0 m/s for 20 seconds, with 2 min intervals of rest in between at 4°C by FastPrep-24 (MP Biomedicals, Solon, OH). The supernatant was collected after centrifugation for 20 min at 4°C and 6500 x g for enzymatic assays. Protein concentrations were determined by a Bradford assay with bovine serum albumin as the protein concentration standard (Bradford, 1976).

Assays of carbonic anhydrase and RuBisCO enzymatic activities

In this study, enzyme activities reported were the average of triplicate assays ± standard deviation. The carbonic anhydrase activity assays were performed as described previously (Gai et al., 2014). The RuBisCO assay is a modified spectrophotometric assay (Lan and Mott, 1991). The activation buffer used as the blank for the spectrophotometer contained 50 mM of bicine at pH 8, 20 mM of $MgCl_2 \cdot 6 H_2O$, and 2mM DTT. The final reaction mixture in the cuvette consisted of activity buffer (5.5% v/v), 10 mM $NaHCO_3$,1 mM ATP, 0.4 mM NADH, 25 units of carbonic anhydrase, 5 units of 3-phosphoglycerate kinase, 5 units of glyceraldehyde-3-phosphate dehydrogenase, 10% v/v cell extract and 0.2 mM of ribulose-1,5-biphosphate (RBP). Everything in this mixture except for RBP was incubated at 30°C for 5 minutes before the absorbance was measured (λ=340 nm) for five minutes. After 100 seconds of measurement in the spectrophotometer, RBP was injected into the sample. NADH concentrations were calculated using an extinction coefficient of 6.22 mM^{-1}.

Quantification of Polymer in R. eutropha

The cell dry weight (CDW) and amount of polyhydroxybutyrate (PHB) was measured as described previously (Karr et al., 1983; York et al., 2003).

Results

Introduction of carboxysomes into R. eutropha strains

Wild type *R. eutropha* H16 has the ability to assimilate CO_2 via the Calvin-Benson-Bassham (CBB) cycle through its possession of operons containing CBB genes (Bowien and Kusian, 2002). However, wild-type *R. eutropha* does not express carboxysomes, and carboxysome shell protein genes are not found in its genome (Pohlman et al., 2006). Thus, we constructed pCarboxy, an arabinose-inducible plasmid with the carboxysome operon from *H. neapolitanus* (Bonacci et al., 2012) and introduced the plasmid to *R. eutropha* strains H16 and Re2061 (ΔphaCAB), which lacks the ability to produce PHB. The use of this plasmid allows the induction of gene expression of carboxysomes when arabinose is added to the growth media. Colony PCR was used to confirm the presence of pCarboxy in *R. eutropha*.

Greater carbonic anhydrase and RuBisCO enzymatic activities in strains with carboxysomes

To investigate whether carboxysomes were being expressed in *R. eutropha* and evaluate the effect of

Figure 3. Enzymatic activity of carbonic anhydrase (A) and RuBisCO (B) in the strain Re2061/pCarboxy, with the strain Re2061 (without pCarboxy) as the control (see Table 1).   Both strains were cultivated in minimal media with 0.05% NH_4Cl with 2% w/v fructose. All flask cultures were induced with 0.2% arabinose (final w/v) after 21 hours of growth. The no arabinose samples were taken 1 hour before arabinose induction, and the post-arabinose samples were taken three hours after arabinose induction. For the carbonic anhydrase enzymatic assay, one Enzyme Activity Unit (U) was defined as the enzyme makes the reaction proceed twice as fast. For the RuBisCO enzymatic assay, the units of enzymatic activity (U) were defined as mmol of NADH consumed/min. Each value represents the mean ± standard error (error bars) of n=3.

carboxysomes on key carbon fixation enzymes, the enzymatic activity of carbonic anhydrase and RuBisCO were assayed in Re2061/pCarboxy, with Re2061 as the control strain (Figure 3). The cultures were grown for 47 hours and induced with 0.2% (w/v) arabinose at 21 hours. For both carbonic anhydrase (CA) and RuBisCO, greater enzymatic activities were seen in Re2061/pCarboxy as compared to Re2061, before and after arabinose induction (Figure 3). The highest CA and RuBisCO activities were seen after arabinose induction; 3 hours after arabinose induction, Re2061/pCarboxy had 11.4 times greater CA enzymatic activity and 2.5 times greater RuBisCO enzymatic activity (Figure 3).

R. eutropha strains with carboxysomes exhibit greater cell growth on formate

To demonstrate the carboxysomes' effect on carbon fixation in *R. eutropha*, we compared the growth of H16/pCarboxy and H16/pBBad when formate was the carbon source. Formate breaks down into carbon dioxide in the media while also providing NADH, effectively providing both carbon and energy sources to *R. eutropha*. The cultures were grown for 65 hours and induced at 20 hours of growth. H16/pBBad, which does not contain the carboxysome operon, was used as the control strain. Initially, the controls strain H16/pBBad grew to a greater optical density, but after eight hours of arabinose induction, H16/pCarboxy began

to equal and slowly overtake H16/pBBad in optical density (Figure 4). After 65 hours of cultivation or 45 hours after arabinose induction, the optical density of H16/pCarboxy was 1.23 times greater than that of H16/pBBad (Figure 4).

R. eutropha wild-type strains with carboxysomes produce less PHB under autotrophic growth

Lastly, to assess the carboxysomes' effect on the production of polyhydroxybutyrate (PHB), a polymer produced by wild-type *R. eutropha*, we measured the amount of PHB produced as a percentage of the cell dry weight (CDW) in H16/pCarboxy with H16/pBBad as a control. Both strains were grown autotrophically for 65 hours with formate as the sole carbon source as described above. At both 48 and 65 hours of growth, H16/pCarboxy was found to produce less PHB than H16/pBBad, with H16/pCarboxy producing no more than 28% PHB and H16/pBBad producing at least 65% PHB (Figure 5). Both strains produced the most PHB at 48 hours of cultivation (Figure 5).

Discussion

Carbon fixation is an important first bottleneck in *R. eutropha's* ability to take in carbon dioxide as a renewable carbon source for valuable products. While *R. eutropha* already has the ability to grow autotrophically on CO_2 (Li et al., 2012), this study aimed to improve carbon fixation in *R. eutropha* by introducing carboxysomes into these microorganism, since carboxysomes have a carbon-concentrating mechanism that makes key carbon fixation enzymes more efficient. After carboxysomes were introduced, *R. eutropha* strain Re2061/pCarboxy exhibited greater enzymatic activity of carbonic anhydrase (CA) and RuBisCO (Figure 3), providing evidence that carboxysomes are being expressed and increase enzymatic activities of the key carbon fixation enzymes,

CA and RuBisCO, in *R. eutropha*. Carboxysomes also led to greater autotrophic growth of wild-type *R. eutropha* strain H16 on formate (Figure 4), implying that carboxysomes improved *R. eutropha's* ability to grow with formate as the sole carbon source. Finally, wild-type strain H16 with carboxysomes was found to produce less PHB than H16/pBBad, the control strain, (Figure 5), implying that an increase in cell growth on formate did not lead to a corresponding increase in polymer formation. The decrease in PHB production may have been caused by competition between carboxysomes and PHB granules for space inside the cell. In current models of PHB biogenesis, small granules coated in protein are continually generated in the cytoplasm before merging together (Beeby et al., 2011). Thus, the assembly of carboxysomes, which are made up of protein shells, could have interfered with PHB granule information. The bacteria may then had channeled more of its carbon flux toward growth, instead of toward PHB production.

Even though PHB production was not increased, this study provides valuable insight into the potential applications of carboxysomes in improving carbon fixation in *R. eutropha*. While many studies have been done characterizing the mechanisms of carbon concentration and assembly in the carboxysome, few studies have been completed that have demonstrated the application of the carboxysome in heterologous hosts. Carboxysomes have been expressed in heterotrophic *E. coli* (Bonacci et al., 2012), and more recently in plant chloroplasts (Lin et al., 2014). However, carboxysomes have not been used to improve carbon fixation and product formation in autotrophic microorganisms. In this study, carboxysomes led to higher enzymatic activities of key carbon fixation enzymes and improved autotrophic growth of *R. eutropha* on formate, thus demonstrating the utility of carboxysomes in improving the efficiency of carbon fixation in an industrially relevant microorganism, while highlighting how carboxysome assembly can be a potential challenge in increasing the yield of products that aggregate in the cytoplasm.

There are many future directions to be explored in how bacterial microcompartments like the carboxysome can be used to improve *R. eutropha's* ability to produce valuable products. Electron microscopy can be used to illuminate how carboxysome protein shell assembly interacts with PHB granule assembly and provide further insight in how carboxysomes assemble in *R. eutropha*. In addition, the introduction of more bicarbonate and CO_2 transporters into R. eutropha may further improve carbon fixation. Recently, overexpressing extra bicarbonate transporters in a cyanobacterial strain led to increased biomass production (Kamennaya et al., 2015), so the addition of bicarbonate transporters in *R. eutropha* may enhance the carbon-concentrating mechanism of the carboxysome. *R. eutropha* has also been engineered to produce branched-chain alcohols as biofuels (Lu et al., 2012), so the expression of carboxysomes into these alcohol-producing strains may lead to an increase in product yield. Since these alcohols are secreted by *R. eutropha* into the medium (Lu et al., 2012), it is more likely that increased carbon flux from improvements in carbon fixation can be directed into greater alcohol production without the challenge of competing against intracellular granules, as with PHB. Thus, there is great potential in utilizing carboxysomes to improve carbon fixation in *R. eutropha* and other bacteria, and bolster their ability to convert carbon dioxide into a variety of products.

Acknowledgments

References

Atsumi, S., Hanai, T., & Liao, J. C. (2008). Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature*, *451*(7174), 86-89. Retrieved on March 30, 2015 from http://www.nature.com/nature/journal/v451/n7174/full/nature06450.html

Beeby, M., Cho, M., Stubbe, J., & Jensen, G. J. (2012). Growth and localization of polyhydroxybutyrate granules in Ralstonia eutropha. *Journal of bacteriology*,*194*(5), 1092-1099. Retrieved on March 30, 2015 from http://jb.asm.org/content/194/5/1092.short

Bertani G., (1951) The mode of phage liberation by lysogenic *Escherichia coli. J Bacteriol* 62:293-300. Retrieved on March 30, 2015 from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC386127/

Bonacci, W., Teng, P. K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P. A., & Savage, D. F. (2012). Modularity of a carbon-fixing protein organelle. *Proceedings of the National Academy of Sciences*, *109*(2), 478-483. Retrieved on March 30, 2015 from http://www.pnas.org/content/109/2/478.short

Bowien B, Kusian B. (2002). Genetics and control of CO_2 assimilation in the chemoautotroph *Ralstonia eutropha. Arch Microbiol* 178:85-93. Retrieved on March 30, 2015 from http://link.springer.com/article/10.1007/s00203-002-0441-3

Bradford M.M., (1976). Rapid and sensitive method for the quantitation of microgram quantiites of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248-254. Retrieved on March 30, 2015 from http://www.sciencedirect.com/science/article/pii/0003269776905273

Budde, C. F., Riedel, S. L., Hübner, F., Risch, S., Popović, M. K., Rha, C., & Sinskey, A. J. (2011). Growth and polyhydroxybutyrate production by Ralstonia eutropha in emulsified plant oil medium. *Applied microbiology and biotechnology*, *89*(5), 1611-1619. Retrieved on March 30, 2015 from http://link.springer.com/article/10.1007/s00253-011-3102-0

Cannon, G. C., Bradburne, C. E., Aldrich, H. C., Baker, S. H., Heinhorst, S., & Shively, J. M. (2001). Microcompartments in prokaryotes: carboxysomes

and related polyhedra. *Applied and environmental microbiology*, *67*(12), 5351-5361. Retrieved on March 30, 2015 from http://aem.asm.org/content/67/12/5351.short

Cannon, G. C., Heinhorst, S., & Kerfeld, C. A. (2010). Carboxysomal carbonic anhydrases: Structure and role in microbial CO_2 fixation. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1804*(2), 382-392. Retrieved on March 30, 2015 from http://www.sciencedirect.com/science/article/pii/S157096390900274X

Ducat, D. C., & Silver, P. A. (2012). Improving carbon fixation pathways. *Current opinion in chemical biology*, *16*(3), 337-344. Retrieved on March 30, 2015 from http://www.sciencedirect.com/science/article/pii/S1367593112000634

Fukui T., Suzuki M., Tsuge T., Nakamura S. (2009) Microbial synthesis of poly((R)-3-hydroxybutyrate-co-3-hydroxypropionate) from unrelated carbon sources by engineered Cupriavidus necator. *Biomacromolecules* 10:700–706. Retrieved on March 22, 2015 from http://pubs.acs.org/doi/pdf/10.1021/bm801391j

Gogerty, D. S., & Bobik, T. A. (2010). Formation of isobutene from 3-hydroxy-3-methylbutyrate by diphosphomevalonate decarboxylase. *Applied and environmental microbiology*, *76*(24), 8004-8010. Retrieved on March 30, 2015 from http://aem.asm.org/content/76/24/8004.short

Kamennaya, N. A., Ahn, S., Park, H., Bartal, R., Sasaki, K. A., Holman, H. Y., & Jansson, C. (2015). Installing extra bicarbonate transporters in the cyanobacterium Synechocystis sp. PCC6803 enhances biomass production. *Metabolic engineering*, *29*, 76-85. Retrieved on March 30, 2015 from http://www.sciencedirect.com/science/article/pii/S1096717615000282

Karr D.B., Waters J.K., Emerich D.W. (1983) Analysis of Poly-β-Hydroxybutyrate in *Rhizobium japonicum* Bacteroids by Ion-Exclusion High-Pressure Liquid Chromatography and UV Detection. *Appl Environ Microbiol* 46:1339-1344. Retrieved on March 30, 2015 from http://aem.asm.org/content/46/6/1339.short

Keys, A. J. (1986). Rubisco: its role in photorespiration. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *313*(1162), 325-336. Retrieved on March 30, 2015 from http://rstb.royalsocietypublishing.org/content/313/1162/325.short

Lan, Y., & Mott, K. A. (1991). Determination of apparent Km values for ribulose 1, 5-bisphosphate carboxylase/oxygenase (Rubisco) activase using the spectrophotometric assay of Rubisco activity. *Plant physiology*, *95*(2), 604-609. Retrieved on March 30, 2015 from http://www.plantphysiol.org/content/95/2/604.short

Lee, S. K., Chou, H., Ham, T. S., Lee, T. S., & Keasling, J. D. (2008). Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Current opinion in biotechnology*, *19*(6), 556-563. Retrieved on March 30, 2015 from http://www.sciencedirect.com/science/article/pii/S0958166908001420

Lee, S. Y., Choi, J. I., & Wong, H. H. (1999). Recent advances in polyhydroxyalkanoate production by bacterial fermentation: mini-review. *International Journal of Biological Macromolecules*, *25*(1), 31-36. Retrieved on March 30, 2015 from http://www.sciencedirect.com/science/article/pii/S0141813099000124

Li, H., Opgenorth P.H., Wernick D.G., Rogers S., Wu T.Y., Higashide W., Malati P., Huo Y.X., Liao J.C. (2012). Integrated electromicrobial conversion of CO2 to

higher alcohols. *Science* 335:1596. Retrieved on March 30, 2015 from http://www.sciencemag.org/content/335/6076/1596.short

Lin, M. T., Occhialini, A., Andralojc, P. J., Devonshire, J., Hines, K. M., Parry, M. A., & Hanson, M. R. (2014). βCarboxysomal proteins assemble into highly organized structures in Nicotiana chloroplasts. *The Plant Journal*, *79*(1), 1-12. Retrieved on March 30, 2015 from http://onlinelibrary.wiley.com/doi/10.1111/tpj.12536/abstract

Lu, J., Brigham, C. J., Gai, C. S., & Sinskey, A. J. (2012). Studies on the production of branched-chain alcohols in engineered Ralstonia eutropha. *Applied microbiology and biotechnology*, *96*(1), 283-297. Retrieved on March 30, 2015 from http://link.springer.com/article/10.1007/s00253-012-4320-9

Pohlmann, A., Fricke, W. F., Reinecke, F., Kusian, B., Liesegang, H., Cramm, R., & Bowien, B. (2006). Genome sequence of the bioplastic-producing "Knallgas" bacterium Ralstonia eutropha H16. *Nature biotechnology*, *24*(10), 1257-1262. Retrieved on March 30, 2015 from http://www.nature.com/nbt/journal/v24/n10/full/nbt1244.html

Price, G. D., & Badger, M. R. (1991). Evidence for the role of carboxysomes in the cyanobacterial CO2-concentrating mechanism. *Canadian journal of botany*, *69*(5), 963-973. Retrieved on March 30, 2015 from http://www.nrcresearchpress.com/doi/abs/10.1139/b91-124#.VRoatPnF8s8

Rae, B. D., Long, B. M., Whitehead, L. F., Förster, B., Badger, M. R., & Price, G. D. (2013). Cyanobacterial Carboxysomes: Microcompartments that Facilitate CO2 Fixation. *Journal of molecular microbiology and biotechnology*, *23*(4-5), 300-307. Retrieved on March 30, 2015 from researchgate.net

Sambrook J., Russell D.W. (2001) Molecular cloning - A laboratory manual, 3rd edition. Gold Spring Harbor Laboratory Press.

Simon, R., Priefer U., Pühler A. (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat Biotechnol* 1(9): 784-791. Retrieved on March 30, 2015 from http://www.nature.com/nbt/journal/v1/n9/abs/nbt1183-784.html

Slater S, Houmiel KL, Tran M, Mitsky TA, Taylor NB, Padgette SR, Gruys KJ (1998) Multiple β-ketothiolases mediate poly(beta-hydroxyalkanoate) copolymer synthesis in *Ralstonia eutropha*. *J Bacteriol* 180:1979-1987.

Yang, Y. H., Brigham, C. J., Budde, C. F., Boccazzi, P., Willis, L. B., Hassan, M. A., ... & Sinskey, A. J. (2010). Optimization of growth media components for polyhydroxyalkanoate (PHA) production from organic acids by Ralstonia eutropha. *Applied microbiology and biotechnology*, *87*(6), 2037-2045. Retrieved on March 30, 2015 from http://link.springer.com/article/10.1007/s00253-010-2699-8

Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C., & Shively, J. M. (2008). Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nature Reviews Microbiology*, *6*(9), 681-691. Retrieved on March 30, 2015 from http://www.nature.com/nrmicro/journal/v6/n9/full/nrmicro1913.htm

York G.M., Lupberger J., Tian J., Lawrence A.G., Stubbe J., Sinskey A.J. (2003) *Ralstonia eutropha* H16 encodes two and possibly three intracellular poly[$_D$-3-hydroxybutyrate] depolymerase genes. *J Bacteriol* 185:3788-3794. Retrieved on March 30, 2015 from http://jb.asm.org/content/185/13/3788.short

▪

On Mortgages and Refinancing

Khizar Qureshi[1], Cheng Su[2]

[1]Student Contributor, Class of 2015, Department of Mathematics, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA

[2]Sloan School of Management, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA

In general, homeowners refinance in response to a decrease in interest rates, as their borrowing costs are lowered. However, it is worth investigating the effects of refinancing after taking the underlying costs into consideration. Here we develop a synthetic mortgage calculator that sufficiently accounts for such costs and the implications on new monthly payments. To confirm the accuracy of the calculator, we simulate the effects of refinancing over 15 and 30-year periods. We then model the effects of refinancing as risk to the issuer of the mortgage, as there is negative duration associated with shifts in the interest rate. Furthermore, we investigate the effects on the swap market as well as the Treasury bond market. We model stochastic interest rates using the Vasicek model.

Introduction

The decision to refinance mortgages is often undermined, and unfortunately, misunderstood by many homeowners. In particular, one overlooked consideration is the closing costs associated with refinancing. Modern real estate finance has [1]developed an intuition for the decision, and subjects it to several factors, including: magnitude of closing costs, change in mortgage rate, borrowing costs, and the remaining time of the mortgage. Interestingly enough, a decrease in mortgage rate or borrowing costs should not always be entertained with the decision to refinance. As we will see, there exists a threshold for the time remaining on the mortgage to validate the net present value of such a decision.

First, we develop a synthetic mortgage calculator with two main features. The first feature is very simple, and conveys to the homeowner the monthly payments associated with a traditional mortgage. The second feature, however, provides the proportions of the mortgage that are allocated towards the interest and principal payments. Upon completion of this mortgage calculator, we will extend its features, and simulate scenarios of refinancing.

The remainder of the paper is organized as follows: an interpretation of the synthetic mortgage calculator includes the time-variant distribution of a monthly mortgage payment to its interest and principal components will be followed by the discrete potential monthly savings available to a homeowner through

1 In Collaboration with Morgan Stanley[2]

2 We thank Professor Peter J. Kempthorne[1]
    for his comments, revision, and guidance

refinancing. Here we will depict the conditions under which an option to refinance should be exercised. Next, we will simulate the growth and breakeven in net present value of a repayment option. We will see that the initial costs are eventually overtaken by the savings in decreased monthly interest payments. The completion of the primal purpose will lead to a discussion of implications on contract duration, bond pricing, MBS hedging as well as short-term interest rates.

Synthetic Mortgage Calculator

Here we present a synthetic mortgage calculator. Given a rate x, a term of N years, a notional of B dollars, we will determine the monthly payment. Further, we will determine the portion of the payment that goes towards paying the principal and the portion of the payment that goes towards paying the interest.

$$B = \sum_{i=1}^{12N} \frac{M}{[1 + \frac{x}{12}]^i}$$

The effective annual interest rate is therefore:

$$EAR = [1 + \frac{x}{12}]^{12} - 1$$

From the monthly payment, M, the interest expense, E_i, is equivalent to:

$$E_i = \frac{x}{12} B_i$$

Where B_i is the outstanding balance. It follows that the remainder of the payment goes towards paying the principal, E_p:

$$E_p = M - \frac{x B_i}{12}$$

Figure 1. The figure above illustrates the portions of the monthly payment that are paid towards interest and principal for a 15-year mortgage with a rate of 3.29 percent. Notice that the portion of the payment allocated towards principal is far greater than that allocated towards interest.

Figure 2. The figure above illustrates the portions of the monthly payment that are paid towards interest and principal for a 30-year mortgage with a rate of 4.03 percent. Notice that initially, the portion of the payment allocated towards the interest is greater than that allocated towards the principal (unlike the 15-year mortgage). However, as time progresses, the payments are weighted towards the principal component.

The proportions, X_i and X_p are:

$$X_p = 1 - \frac{xB_i}{12M} \text{ and } X_i = \frac{xB_i}{12M}.$$

Refinancing Cash Flows with Prepayments

Let S = the monthly prepayment rate for n months. Also suppose we have the following variables in month n without prepayment risk: Balance, B_n, Interest, I_n, and Principal, P_n. The total monthly payment in month n is simply:

$$\hat{M}_n = \frac{\hat{B}_n i (1+i)^{N-n+1}}{(1+i)^{N-n+1} - 1} = M_n Q_{n-1}$$

with a scheduled principal portion of the monthly payment equivalent to:

$$\hat{P}_n = \frac{\hat{B}_n i}{(1+i)^{N-n+1} - 1} = P_n Q_{n-1}$$

and an interest portion equivalent to:

$$\hat{I}_n = \hat{B}_n i = I_n Q_{n-1}$$

and unscheduled principal payment in month n equivalent to:

$$P_n^* = S(\hat{B}_{n-1} - \hat{P}_n)$$

and a remaining balance at the end of the month equivalent to:

$$\hat{B}_n = B_{n-1} - \hat{P}_n - P_n^* = B_n Q_n$$

Through refinancing, the interest savings over time is simply:

$$\sum_{i=n}^{N} \hat{I}_i - I_i.$$

It follows that for longer maturities, the accrued interest savings on a mortgage are higher.

Savings

We now consider the option of refinancing such a mortgage. Namely, given mortgages A and B, and a refinancing cost C, should mortgage A be refinanced to mortgage B for immediate cost C? And if it should, what is the time required to break even? We will see that such a decision is dependent on factors such as: current interest rate, new interest rate, closing costs, and the time planned to remain in the home.

Given a new interest rate, X_{new}, the monthly payment is:

$$M_{new} = B \frac{x_{new}}{12} \left[1 - \frac{1}{\left[1 + \frac{x_{new}}{12}\right]^{12N}} \right]$$

It follows that the monthly savings, S_i, is simply:

$$\hat{I}_i - I_i.$$

The mortgage should only be refinanced if the net present value of doing so is positive. That is:

$$NPV = -C + \sum_{i=n}^{N} [\hat{I}_n - I_n] \geq 0$$

Therefore, the number of months required for the interest and principal mortgage insurance savings, S to exceed closing costs, C are:

$$t_{Months} = \frac{C}{S_i}$$

However, there exist a few situations in which refinancing should not be done, even with lowered interest rates:

- Only a short term structure for the loan remains
- Prepayment penalties exceed expected savings
- Mortgage rates are expected to drop further

Simulations

To develop a practical understanding of refinancing, we simulate a specific example. Suppose there is a 20-year outstanding mortgage with an initial loan amount of 100,000. Further, assume that there is a state of nature in which there is a 1 percent drop in interest rates. The homeowner is tempted to borrow at this lower interest rate, and prepay, but should be wary of closing costs. In particular, the homeowner should ponder the following: Given the closing costs, and the remaining time period on the mortgage, is the net present value of refinancing positive? For any rational monetary decision, this is a strict condition. Below is a figure that illustrates the results of the simulation. Notice that due to closing costs, the initial $NPV_{t=0}$, is negative. However, 15 months into the refinancing decision, the NPV breaks even.

Consider the very same mortgage, but with an outstanding term of less than 15 months. It follows that with the same closing costs and base mortgage rate, the net present value of refinancing will surely be negative. That is, t<15years, NPV_T <0. This is because the payment structure is not allotted sufficient time to break even the costs. In scenarios such as these, the mortgage should never be refinanced.

Next, we generalize our simulations to illustrate the difference in activity for a 15-year and 30-year mortgage. In particular, with and without closing costs, we try to determine the effects of refinancing on the payment structure for two traditional contract lengths. First, we consider an unrealistic case in which rates are lowered, and no closing fees are imposed:

The simulations above aid in the development of an intuition for mortgage activity for the following states:

- Refinancing without closing costs (15-year mortgage)
- Refinancing without closing costs (30-year mortgage)
- Refinancing with closing costs, λ (15-year mortgage)
- Refinancing with closing costs, λ (30-year mortgage)

It is clear that closing costs, depending on magnitude may significantly reduce the perceived benefits of refinancing through means of increasing the effective monthly payments.

Figure 3. The figure above illustrates the number of months [15] required for the net present value to breakeven during refinancing. Suppose there is a 20-year (outstanding) mortgage with an initial loan of 100,000. Although there is a 1 percent drop in interest rates, there is a remaining balance of 85812, and closing costs are approximately 1000. The new loan issued through refinancing is simply the sum of the remaining balance and the quoted closing costs.

Figure 4. The graph above illustrates the reduction in monthly coupon with refinancing. Here we assume zero costs to refinance and present the effects in both a 15-year and 30-year mortgage. Notice that because of the unrealistic zero costs, the effective monthly payment is reduced significantly. However, as time progresses, the monthly payment converges. This holds for both a 15-year and 30-year mortgage.

Figure 5. The graph above illustrates the reduction in monthly coupon with refinancing. Here we assume fixed, non-zero costs to refinance and present the effects in both a 15-year and 30-year mortgage. In particular, to account for the costs, the 15-year and 30-year mortgage rates are given a multiplier λ =1.20. Effectively, this increases the capital commitment to replicate closing costs. Notice that due to the newly introduced costs, the effective difference in monthly payments has narrowed significantly. It is also observable that the difference narrows more slowly with the 30-year mortgage, than it does with the 15-year mortgage. and the quoted closing costs.

Negative Duration Risk

In most mortgages borrowers are given prepayment options, and are typically exercised when interest rates fall. Homeowners are able to borrow at these lower interest rates, and in most cases, effectively reduce their monthly payments. However, this results in a negative convexity for Mortgage-Backed Securities (MBS). As interest rates fall, the duration for Mortgage-Backed Securities actually shortens. Inversely, an increase in interest rates discourages homeowners from borrowing and refinancing, resulting in a positive duration. The consequences of such an effect are quite significant, and may be readily observed. The equation below models the effect on duration, D, with a change in yield (interest rates), y (Myers, 2011).

$$D = -\frac{dP}{dY}\frac{1+y}{P}$$

Notice that with an increase in interest rates, duration typically decreases. However, in the case of MBS, duration decreases due to prepayment and refinancing effects.

As a response to the shortening duration, market participants hedge with interest rate swaps, and opt to receive the fixed leg, or equivalently, terminate any floating receivables. Moreover, such an effect can result in buying pressure on the "swap bond", which in turn results in temporary decrease in the swap spread. The mechanism above is opposite of traditional bond activity, justifying the compensation market participants demand for negative duration risk. This level of intuition may be valuable when judging credit market activity.

It is critical to not only understand the effects of duration on the credit market, but also how interest rates vary over time. In particular, we explore the Vasicek model, which correctly hypothesizes the

stochastic movement of interest rates. We then extend the model parameters to bond pricing, which is of interest to market participants who seek to proliferate from interest rate movement. Finally, we return to the idea of hedging, and why the MBS swap market is so liquid.

Vasicek Model for Interest Rates

The Vasicek Model in a continuous-time term-structure model for short rates. It follows a stochastic mean reversion process as follows (Vasicek, 1977):

$$dr_t = \kappa(\bar{r} - r_t)dr + \sigma dW_t$$

where r_t is the short rate, κ the mean reversion coefficient, and dW captures a Wiener Process (Vasicek, 1977).

Implication on Bond Pricing

The price of a T-year Zero Coupon Bond may be determined by:

$$P = e^{\alpha + r\beta}$$

$$\beta = \frac{e^{-\kappa T}-1}{\kappa} \text{ and } \alpha = \bar{r}\left[-\beta - T\right] + \frac{\sigma^2}{2\kappa^2}\left[\frac{1-e^{-2\kappa T}}{2\kappa} + \frac{2\beta}{\kappa} + T\right]$$

Once κ is determined from the Vasicek model, it may be used to determine the resulting bond price. It is readily observable that a larger value of κ, or equivalently, a faster mean reversion process implies lower bond prices. This increase in yield translates to a narrowing in swap spreads.

Implication on Short Rate

Consider a three-month T-bill with rate r_t at time

t. The vasicek model, in discrete form, implies that at time t+1, the T-bill will carry rate:

$$r_{t+1} = \kappa\bar{r}\Delta + (1 - \kappa\Delta)r_t + \sigma\Delta^{1/2}\epsilon_{t+1}$$

where σ is the one-month volatility, Δ a change in time of 1/12 years, and ϵ a small shock such that $\epsilon \sim N(0,1)$. Note that in practice, the short rate would follow a multi-factor model.

Implication on MBS Hedging

Consider a swap spread λ-T where λ is the swap rate for a mortgage-backed security. It follows that with in increase in short rate rate, the swap spread would decrease. This is attributable to two factors. One straightforward reason is that the bond yield would increase in response to the increase in short rate. The second reason, however, is that with an increase in duration, market participants will choose to sell duration through means of paying the fixed leg of the swap (or equivalently, sell current receivables), which effectively lowers the swap rate. A decrease in the swap rate and an increase in bond yield result in drastic decreases in the swap spread. Inversely, decreases in interest rates result in large increases in the swap spread.

Conclusion

A decrease in interest rates provides incentive to refinance mortgages. However, as we have seen, an important consideration is the savings on the mortgage taking into account the costs for refinancing. Only if the 15-year or 30-year sum of the change in monthly payments exceed the borrowing costs should the mortgage be refinanced. In particular:

- Differential changes in the interest rates should not be reasons to refinance.
- There exists a minimum time T>0 for the remaining life of the mortgage only under which it is optimal to refinance
- Closing costs are almost always non-zero, so the preceeding conditions are therefore strict.

Refinancing and prepayment also provide sources of risk to mortgage issuers in the form of negative duration. Issuers face the risk of having the sum of the time-weighted payments reduced. Compensation to carry such risks are likely factored into the mortgage rate offered to the homeowner depending on his/her individual propensity to refinance or issue prepayments.

Changes in the interest rate have implications on the MBS market, treasury bond market, and swap market. A decrease in the interest rate can lower the prices of mortgage-backed securities, decrease the treasury bond yield, and widen the swap spread. Inversely, increases in interest rates can increase the price of MBS, increase the treasury bond yield, and narrow the swap spread. Such movements can be modeled through market participants response in buying/selling fixed legs of MBS swaps.

Future Considerations

In future studies, we wish to test the feasibility of the Vasicek model for term structure. In particular, we wish to determine the effectiveness in determining the change in swap spread due to negative duration hedging. We suspect that the model is insufficient for capturing all sources of market risk, and that macroeconomic, systemic and credit factors will be necessary. In modern practice, market participants use modified versions of the Heath-Jarrow-Morton (HJM) model or the multi-factor Hull-White model. There is also the unexplored possibility of using a modified version of the Cox-Ingersoll-Ross stochastic model for interest rates. To test the effectiveness of the model, 6-month, 12-month, 5-year, 15-year, and 30-year term structures will be used. Both the credit markets and mortgage markets are fascinating, and it remains to be seen the effects of efficient refinancing on the asset pricing of underlying instruments.

References

[1] Oldrich Vasicek, An Equilibrium Characterisation of the Term Structure,. Journal of Financial Economics, 1977.

[2] Brealey, Richard A.; Myers, Stewart C.; Allen, Franklin, Principles of Corporate Finance New York, NY: McGraw-Hill Irwin, pp. 50-53

MURJ 2015 CPW Research Poster Competition

By Pratheek Nagaraj

For the second year in a row, MURJ has hosted a competition for undergraduate students from across all majors to present their research to the wide MIT community. This interdisciplinary competition takes place during Campus Preview Weekend (CPW), a time where admitted high-school students, colloquially termed pre-frosh, visit campus as part of their decision making to enroll as undergrads at MIT. Because of this overlap the turnout to the competition is varied and diverse with attendees ranging from pre-frosh and their family to graduate students and post-docs. This year the competition featured work from seven students across six different majors including Biological Engineering, Mathematics, and Brain and Cognitive Sciences.

Communicating research to diverse audiences is a critical skill that is important to the research process. MURJ believes that this competition provides an excellent platform for many students to develop and hone this important skill and so we encourage MIT undergraduates doing research to participate next year.

This event would not have been possible without the support of many. First, MURJ would like to acknowledge the support from the Biology and Mathematics departments for generously providing food and funding for awards respectively. Second, thank you to the judges who kindly donated their time to engage with undergraduate research and give feedback to the presenters. And third, MURJ appreciates the support from the Baker Foundation for its assistance with equipment.

The MURJ team is excited to make this event even better for next year and look forward to your participating or attending! Thank you to all the presenters who shared their exciting research with the MIT community. Congratulations to our winners, Nelson Hall '16 – Overall Winner, Daphne Superville '17 – Academic Content Winner, and Kara Presbrey '16 – Presentation Winner.

MURJ poster session, CPW 2015. Photo: Pratheek Nagaraj

One Focus: committed to improving the lives of cancer patients everywhere

We are excited to announce the launch of **TAKEDA ONCOLOGY** formerly known as **MILLENNIUM: THE TAKEDA ONCOLOGY COMPANY.** Our mission is unchanged; we endeavor to deliver extraordinary medicines to patients with cancer worldwide through our commitment to science, breakthrough innovation and passion for improving lives.

By concentrating the power of leading scientific minds and the vast resources of a global pharmaceutical company, we are finding innovative ways to improve the treatment of cancer. While we've made great strides in our fight against cancer, we are determined to do more — to work harder and to achieve greater. Our singular focus drives our aspirations with the same passion, agility and entrepreneurial spirit that has always been at the heart of our culture and made us the leaders in oncology that we are today.

Takeda Oncology, we aspire to cure cancer.

To learn more, visit us at **takedaoncology.com.**

©2015 Millennium Pharmaceuticals, Inc. All rights reserved.