Dear Reader,

There is no end to scientific discovery, with deeper analysis improving our understanding in various fields in science, technology, engineering, and math at an ever-increasing rate. From biology to computer science, researchers continue to find inspiration both in their own fields and in those of others, prompting them to look deeper for new answers and enlightenment. Acknowledging the depths to which science takes those willing to explore what it has to offer, we take a deeper look in issue of the DUJS.

The staff review articles for this issue cover a variety of disciplines and topics. Hyungdon Joo explores the emerging field of host-guest chemistry and its uses in materials science, medicine, and engineering. Brenda Miao takes a look at major depressive disorder, discussing the leading theories and treatments in the field. Krishan Canzius provides a deeper look into fractals, providing both the mathematical basis and history of these patterns.

Next, Cara Van Uden delves into deep learning, evaluating the current leaders in the field, as well as the complications that accompany training neural networks. Sam Reed discusses the astounding advances that have been made in sequencing since the Human Genome Project first sequenced the human genome in 2001.

Taking a deeper look at life, Sam Greydanus explains how the brain uses visual cues to perceive depth. Christine Park evaluates both bottom-up and top-down approaches in systems biology, both of which address the complexity of biological systems in developing new disease treatments. To complete our review articles, Kevin Kang evaluates deep vein thrombosis, explaining its prevalence in professional athletes.

Our winter issue features a faculty interview with Devin Balkcom, Associate Professor of Computer Science, where he discusses the connection between computer science and robotics, as well as his current research. The first and second place winners of our fourth annual International Science Essay Competition for High School Students, Emily Xu and William (Woosung) Jung, are also featured in this issue for their winning essays.

We have original research submissions from two Dartmouth undergraduates: Arvind Suresh, detailing his research on angiogenesis, and Sarah Bessen, discussing the potential of the Mediterranean diet in gastroparesis treatment. Finally, we are proud to support four Dartmouth alumni, three of whom are DUJS alumni, in announcing the release of their new book *What Every Science Student Should Know*, a guide for students interested in pursuing STEM.

We would like to thank our writers, editors, staff members, and faculty advisors for making this issue of the DUJS possible. Without the support of the Dartmouth community, we would not be able to maintain our success as an outstanding scientific outlet.

Thank you for reading the DUJS, and we hope that you enjoy the issue.

Sincerely,

Stephanie Alden
*Editor-in-Chief*
## Table of Contents

**Faculty Spotlight: Devin Balkcom**  
*Nan Hu ’18*  

**Molecular Caging: Trapping Molecules**  
*Hyungdon Joo ’19*  

**New Theories of Depression**  
*Brenda Miao ’19*  

**Zooming Deeper into Fractals**  
*Krishan Canzius ’18*  

**Deep Learning: Can Computers Learn Like We Do?**  
*Cara Van Uden ’19*  

**Deep Sequencing: The Rapid Genomics Revolution**  
*Sam Reed ’19*  

**Depth Perception: More Than Meets the Eye**  
*Sam Greydanus ’17*  

**Systems Biology: Drug Discovery at a Deeper Level**  
*Christine Park ’17*  

**Deep Vein Thrombosis and its Incidence in Athletes**  
*Kevin Kang ’18*  

### ISEC WINNERS

**David and Goliath: The Future of Targeted Therapy in Cancer Treatment**  
*Emily Xu, Raleigh Charter High School*  

**Bioinformatics: Piercing and Pinpointing Moonlighting Proteins and Beyond**  
*William [Woosung] Jung, Korea International School*  

### BOOK ANNOUNCEMENT

**What Every Science Student Should Know**  
*Yoo Jung Kim ’14 and Andrew H. Zureick ’13*  

### ORIGINAL RESEARCH SUBMISSIONS

**Evaluating Pathological Angiogenesis in Wild Type vs. Slug Knockout Mice**  
*Arvind Suresh ’19*  

**The Effects of the Mediterranean Diet on Symptomatic Patients with Gastroparesis**  
*Sarah Bessen ’16*
Not many professors at Dartmouth actively do research in the field of robotics. What about robotics really appeals to you, and what makes you interested in robotics?

Robotics is not necessarily what I thought it to be when I went to graduate school. I had not done any robotics when I went to study for a Ph.D. in robotics. I thought it was humanoid robotics, programming those robots. It turns out that there are lots of problems that we call “robotics” because they are cool and because we are roboticists, but it is a really nice mix of a lot of different areas. You get to do some physics, some design and building, and some programming stuff. You get to do some formal analysis of mathematics and some really rigorous work. I really like large combinations of things. One of the most exciting things about it, subject wise, is that you get to program the physical world. I like computer science and programming, and I like to see things move, and we get to do that.

Robotics covers many different areas. What do you think makes robotics robotics? What common elements tie these areas together?

That is actually hard to say. In some sense it is the integration. We typically have some software, some analysis, some physics, and some moving device in the world. This is also found in a dishwasher or a car, but we wouldn’t necessarily call them robots, so there’s also this sense of being on the envelope of what we can do. A lot projects I do actually don’t necessarily have direct control of the robot through software, so sometimes we’re interested in a specific component or a special purpose robot for a special purpose task. That is really exciting too, and that is still robotics.

One area of research you explore is laundry and origami folding. Do you think that we will ever be freed from folding our own laundry? What are some of the challenges you see?

Well, I think we will. Is it five, 10, 50, or 100 years? I do not know. It’s hard to predict those things, but the capability to develop that technology is definitely possible. Sometimes you have to
actually modify the world to make these things work. We can
definitely build a laundry-folding robot if everyone is going
to wear the same uniform, with devices [attached] to help for
folding. The question is really the flexibility of [the machine].
Will it fold some of your laundry or all of your laundry? Does
it fold the towels? Does it fold the t-shirts? The challenges
are really that adaptability to new situations. It turns out that
things like folding laundry are quite easy if the laundry is flat,
even if it is of a somewhat unpredictable shape. The flattening
is harder, but I think it is possible as well.

You have already mentioned that robotics is a field
that allows you to do a lot. What advice do you have for
someone who would like to go into robotics?

If there are lots of different areas, one approach is to become
technically proficient in all of those different areas. But
sometimes, actually, robotics motivates a person to get better
at something. I was quite a capable programmer when I went to
graduate school, and that was sufficient for me to contribute to
various teams and projects. My mathematics skills were much
more limited. I had enjoyed mathematics, but I had not been
really motivated by it. Robotics actually motivated my interest
in mathematics quite a bit. Now, a lot of my papers are largely
mathematical in nature. I think just get[ting] involved is the
first answer. Pick whatever skill you find most satisfactory and
become proficient in that; then you should have an open mind
when learning things and getting involved. What is so good
about robotics is that it attracts people who do not narrowly
define themselves.

You have taught a variety of classes at Dartmouth, from
introductory computer science to advanced courses like
robotics and artificial intelligence. Which has been your
favorite to teach and why?

That is hard to answer because, of course, they all have a
different flavor. I love teaching the upper level courses because
those students are really excited about what I do, they get
involved in the lab, and they do cool projects. But I also really
like teaching the big introductory CS1 class because it is a
different stage. I remember learning some of that material—
some of it on my own, some of it in classes. It is just really
exciting to be able to share that with people for the first time.

In addition to teaching Dartmouth classes, you run
a robotics summer camp for middle and high school
students. Do you think that it is valuable to have
exposure to computer science early on?

I do. I should also tell you that I have not been running the
camp the last couple of years because I had other things going
on, but it is very valuable. I use it as a vehicle a little more
for teaching a bit of basic programming. I think it is really
important that students get a real, concrete skill. There are
a lot of camps that are good, but it is not really clear what
concrete skill was learned. I think if you give someone a little
bit of programming and show them that they are really good
at it, then they can go further on their own.

Why specifically robotics? What are some projects that
the campers have worked on?

Well, because I have robots, and it is a hook. If [the students]
want to program robots, that is great. Some projects include
programming the robot arm to draw pictures. They have
programmed vacuum robots to follow paths and to move
around and explore a maze. They have programmed the robot
arms to build things with Legos. For the middle school camp,
we have used Lego-style robots, and they have programmed
them to do wall-following and to coordinate to put on a dance
performance.

In the first artificial intelligence (AI) class this term
you mentioned that AI is a subfield of robotics. Can you
explain what you mean by this?

Well, I said that, as a roboticist, I prefer to believe that everything
is a subfield of robotics, including artificial intelligence. You
get a robot, and you are going to need artificial intelligence.
I think an artificial intelligence researcher would like to view
it the other way around, though. Artificial intelligence is
about making decisions automatically with software, using
algorithms for doing that, and modelling the world. Many
robots need to do that, whether directly or indirectly.

On the topic of humanoid robots, do you think that
making a robot that acts human is more of a hardware
or a software challenge?

I think it is largely a software challenge. Obviously the
hardware has a long way to go. People in artificial intelligence
have been making wild claims for the past 60 years and more,
like "we will have human-level intelligence within 10 years," and
every 10 years someone will come along and say it will
just be another 10 years. I think if they were at a different
time scale, like 10,000 years, maybe it would make more sense.
Once you get into artificial intelligence, you realize we are not
even close. We do not understand how the brain works; we
do not understand how to simulate it. We can make decisions
and solve certain types of problems, and we can beat humans
in certain specific domains such as chess. But the idea of a
"singularity" within 10 or 15 years, I do not think that is
possible.
Fantastic Voyage

Ever since Democritus introduced the concept of the atom to mankind, scientists and laymen alike have been fascinated by exploring deeper into the world of the miniscule. In the movie Fantastic Voyage, a team of surgeons and a submarine shrink down to a size that allows them to enter a patient’s bloodstream and travel to a cancer site, where they perform non-invasive, precision surgery on a tumor. Similarly, in the Marvel comic book series Ant-Man, Dr. Henry Pym discovers a serum capable of reducing molecular distances, shrinking any object into the size-scale of an ant. Throughout the years, scientists have experimented with producing similar feats; in this case, however, they place smaller molecules within larger ones. While these teams of chemists, physicists, and biologists have yet to recreate a novel shrinking technology, their work encompasses the heart of these movies and wild dreams—the manipulation of technology and medicine at the nanoscopic level. With advancements in chemical synthesis, molecular imaging, tumor identification, and targeting techniques, molecular caging has become a nascent reality in nanomedicine and nanotechnology. This article will provide a summary of the history of molecular caging, its present applications in medicine and technology, and its technological implications.

The Origins of Molecular Cages

Crown Ethers and Host-Guest Chemistry

Molecular cages were originally conceived in the field of host-guest chemistry (1, 2, 3). In host-guest chemistry, a relatively large host molecule, usually composed of a cage or a ring, contains or binds reversibly to a smaller guest molecule via non-covalent interactions, such as hydrogen bonds or dipole-dipole interactions (1, 2, 3). Host-guest chemistry has existed since scientists first noticed molecules contained within other molecules in residues from their experiments.

The field was not formally recognized until Charles J. Pedersen, Donald J. Cram, and Jean-Marie Lehn received the 1987 Nobel Prize in Chemistry (1). In the 1960s, Charles J. Pedersen sparked to life the study of host-guest chemistry with his discovery of the crown ether, inspiring both Cram and Lehn to direct their research toward this field (3, 4). In his Nobel Lecture, Pedersen elaborated upon the process by which he synthesized, analyzed, and concluded that the crown ether existed (2). Thus, the crown ether was the first identified host molecule. It complexed with alkali cations and, in fact, required the presence of alkali cations during synthesis to facilitate the completion of its ring structure (2).

Lehn and Cram followed in the wake of Pedersen’s discovery, each adopting different approaches to furthering the field. Lehn focused his own work on the exploration of another class of host molecules called cryptands. He directed his work toward exploring the mechanisms by which host molecules recognized compatible guest molecules (3). By applying his work on molecular recognition toward the analysis and synthesis of receptors and molecular switches, Lehn expanded host-guest chemistry into the current field of supramolecular chemistry, which inspired biologists and nanomaterials researchers alike (3). Meanwhile, Cram was lauded for his systematic development of Corey-Pauling-Koltun (CPK) molecular models of potential host molecules and his subsequent synthesis of models to confirm the properties of the most potentially favorable structures (4). Together, these three men represent the pillars of molecular caging and the birth of an entire field of academic research, which led to yet another milestone discovery in the 1990s.

Figure 1: A cucurbituril molecule host-guest caging a molecule that can suppress cancer tumors.
Second Generation Host-Guest: Cyclodextrin

Following Pederson’s ground breaking discovery, chemists continued to expand host-guest chemistry, sifting through existing molecules for host-guest pairings, until they uncovered cyclodextrin—a ring shaped host similar to crown ethers (5). Cyclodextrins are abundant in nature, meaning they do not require extensive synthesis reactions (5, 6). Unlike crown ethers, cyclodextrins consist of two rims and a cavity that have opposite polarity, with polar hydroxyl groups on the rims and nonpolar functional groups within the cavity (6). Among the pioneers researching cyclodextrins, Ronald Breslow began the movement toward mimicking biological enzymes with host-guest macrocycles—cyclic molecules with host-guest properties (5, 6). In particular, Breslow exploited cyclodextrins’ binding affinity for esters such as cocaine and cholesterol to synthesize promising, medically applicable catalysts (6). Cyclodextrins marked a new generation of host-guest molecules, introducing catalytic mimicry to host-guest chemistry.

The Present and Third Generation

Host-guest chemistry was founded with the discovery of the crown ether and furthered by cyclodextrin development. Then, through the efforts of David Reindouht, Tomoki Ogoshi, and Kim Ki Moon, host-guest chemistry entered the third and present generation of host molecules consisting of calixarenes, pillararenes, and cucurbiturils, respectively (7, 8, 9). Calixarenes and pillararenes possess similar structures, composed of fused benzene rings containing additional functional attachments that alter their chemical properties (7, 8). Calixarenes, however, are shaped more like baskets, with polar hydroxyl groups on their lower rim and non-polar functional groups on their upper rim and cavity (7). Meanwhile, pillararenes, which were discovered in the NMR spectra of Tomoki Ogoshi, are shaped more like columns (8). While pillararenes and calixarenes are similar in structure, they exhibit distinct functions and properties (8). Calixarenes are very versatile in their synthetic, catalytic, and supramolecular behaviors, but they are highly susceptible to perturbation in acidic solutions (7). Pillararenes provide a close alternative, retaining comparable host-guest binding stability under conditions where calixarenes fail (7, 8). The latest of the third generation host-guest molecules to be discovered, cucurbiturils, are “pumpkin shaped” columns of urea, with various caging properties of particular use in transport of medicine throughout the body (9). The identification of these three molecules enabled biological mimicry and organic catalysis to form from host-guest chemistry, introducing novel molecular medical techniques and catalytic mechanisms.

Applications

Nanomedicine

Among the first fields to use molecular caging was medicine. It was immediately apparent that host-guest chemistry could be applied to the injection of drugs, genes, receptors, and other biologically functional molecules within capsules for transportation, targeting, and protection (8).

Two other promising supramolecular applications for medicine include cyclodextrins and molecular “tweezer” molecules. Like crown ethers, cyclodextrins prevent drugs from releasing prematurely and facilitate drug transport, especially those of a hydrophobic or overly reactive nature, to target tissue (1). Meanwhile, molecular “tweezers” consist of two open arms conjoined at an anchor point. A cavity that selectively recognizes, binds, and removes target molecules lies between these arms (1). Recently, this technology was applied to remove sterol from dairy products and other food sources to reduce harmful cholesterol intake by humans (1). These applications mark advances toward less invasive and more specific treatment of diseases, as well as disease prevention via molecular treatment of foods and other substances.
Manipulation of Cells, Proteins, and DNA

In addition to their contributions to medicine, molecular cages have provided biologists with a vast array of tools for the manipulation of cells and the development of nanotechnology. For biologists, photo-induced molecular cages provide the most promising advances in manipulating the transport of proteins, DNA, neurotransmitters, and other cellular materials. These molecules are released when their surrounding molecular cage is exposed to light (11). Ellis-Davies (2007) lists the methods of synthesis and analysis of efficiencies of molecular cages formed with various guest molecules ranging from ATP and secondary messengers to DNA and protein (11). Photochromic molecules, such as ortho-nitrobenzyl protecting groups, reversibly bind and render trapped molecules biologically inert (11). The photoremovable protecting molecules are versatile in their ability to trap biologically functional materials, providing efficient means for transporting secondary messengers, neurotransmitters, and even entire gene sequences past cellular membrane barriers (11). Finally, titanium:sapphire lasers can illuminate these photo-inducible molecular cages to stimulate payload release (11). Biologists in particular may find the ability to tune the location, magnitude, and frequency of the payload release of these photoremovable molecular cages useful (11). This control gives researchers unprecedented levels of precision during cell manipulation (11).

Meanwhile, some biologists have opted to create cages out of the very payloads themselves, composing structures out of DNA, RNA, and protein (12). In a groundbreaking synthesis reaction, Lai et al. (2014) successfully fused oligomeric proteins with alpha-helix linkers to form a cube with the largest pore size recorded to date (12). The size of these central pores allows protein fusion cubes and other structural DNA molecular cages to serve both as vehicles for nanomedicine and as biological reaction chambers, potentially making the cell itself the laboratory of future microscopic biochemical experiments (12).

Nanomaterials

Molecular cages and supramolecular chemistry have also contributed significantly to the development of novel nanomaterials for catalysis and filtration. One such example is zeolite, a metal-organic framework (MOF) composed of aluminosilicate that has various exploitable supramolecular properties (1). Zeolite consists of a supercage and porous framework, both of which allow for selective filtration of materials and incorporation of catalytic atoms into its structure (1). More specifically, zeolites are perfect for hosting transition metals. These zeolite-metal combinations act as potent heterogeneous catalysts that benefit from facilitated reagent/catalyst separation and a non-volatile, non-reactive zeolite backbone. Other metal-organic frameworks have potential as containers for storage; for example, scandium-based MOFs have been used to store hydrogen gas fuel (10). By combining scandium-based metal-organic frameworks with binuclear μ-OH bridges, Ibarra et al. (2011) gave MOFs permanently porous structures, which provide the perfect platform for reversible gas storage, the essence of fuel tanks (10).

The Future

With the development of applications in cancer treatment and targeted drug delivery, molecular cages hold promise for producing less invasive therapies and improving drug performance by targeting treatment to relevant target tissues (1, 6, 7, 8, 11, 12). Furthermore, molecular cages may advance targeted genomic
transcription, incorporation of genetic sequences, and gene preservation (8, 11). However, such advances in technology come with weighty implications concerning the ethics and efficiency of developing precision drug and gene delivery methods. Questions arise concerning the immunity of individuals subjected to precision drug therapies. While research indicates that cancer patients display far healthier and more effective antitumor responses to targeted treatments, further research is required to expand targeted treatment to the delivery of painkillers and other medications (7, 8).

Meanwhile, chemists strive to improve methods of synthesis, separation, and exploitation of supramolecular structures. Dartmouth Professor Chenfeng Ke recently proposed a new deprotection-followed-by-activation strategy for synthesizing a class of supramolecular structures called pillar[5]arenes. Pillar[5]arenes, named according to the number of repeating molecules that make up their diameter, share many properties with cyclodextrins and crown ethers (13). In his 2015 paper, Professor Ke describes the failure of conventional silica gel chromatography to completely separate all regioisomers of pillar[5]arene and provides a detailed explanation of his alternate separation method (13). This method both improves synthesis yields and facilitates the understanding of macrocycle isomerization (13). Other researchers, such as Carlos-Andres Palma of the Technische Universitat Munchen, are exploring the thermodynamics of gas molecules by trapping them within predictable trajectories in two-dimensional molecular cages (14). This novel approach to observing the thermodynamics of materials may improve researchers’ ability to predict the properties of stored gases and those of other molecules (14).

Molecular caging presents a venue for deeper insight into fields beyond its birthplace in chemistry. Nanoscale development in the fields of medicine, materials science, and engineering demonstrates molecular caging’s potential for use in novel treatments and nanoscale technologies. By simply placing smaller molecules within larger molecules, researchers have brought a new depth of knowledge to the microscopic world.

“Nanoscale development in the fields of medicine, materials science, and engineering demonstrates molecular caging’s potential for use in novel treatments and nanoscale technologies.”

References

An Overview of Major Depression

Centuries ago, the people of Ancient Greece believed that depression, which they called “melancholia,” was caused by an imbalance in the four humors, or bodily fluids (1). Today, scientists possess a much greater understanding of the biological and environmental causes of this disorder, which range from genetic predisposition to substance abuse (2). Despite these advances, depression still cripples the health of millions, so researchers continue to develop new treatments.

One form of depression currently being studied is major depressive disorder (MDD), a severe and often recurrent form of depression. The Diagnostic and Statistical Manual of Mental Disorders (DSM), which provides the standard criteria for diagnosing mental illnesses, states that those with MDD must exhibit at least 5 of 9 symptoms. These symptoms include: 1) depressed mood; 2) inability to feel pleasure, or anhedonia; 3) changes in weight or appetite; 4) changes in sleep habits; 5) anxiety or retardation; 6) fatigue; 7) feelings of guilt or worthlessness; 8) decreased concentration; and 9) recurrent thoughts of death or suicide (3).

MDD is the costliest mental illness worldwide, both in terms of quality of life and monetary loss (4). In the US alone, an estimated 33 million adults, about 16.2 percent of the population, will suffer from a major depressive episode at least once in their lifetime (4). Approximately 60 percent of people diagnosed with MDD report an inability to live independently (4). The economic burden of MDD in the United States, including health care expenditures and costs to employers for lost workdays, totals about $30 billion per year (2).

Researchers have spent decades researching and developing treatments in an attempt to reduce the severity and prevalence of MDD. However, the most common treatments, such as monoamine antidepressants, are not always effective and may take several weeks to produce their desired effects (5). As a result, experts are conducting further research to discover new treatments to reduce the burden of MDD.

The Monoamine Hypothesis

The predominant theory for MDD is the monoamine hypothesis, developed in the 1960s by scientists studying monoamine oxidase inhibitors (6). Monoamine neurotransmitters, such as serotonin, norepinephrine, and dopamine, have been found to alleviate depressive symptoms. Monoamine oxidases removes these chemicals from the synaptic cleft, terminating their effects (7). As their name implies, monoamine oxidase inhibitors block monoamine oxidase activity to prevent monoamine degradation, prolonging their influence. Because monoamine neurotransmitters lessen depressive symptoms, researchers hypothesize that low concentrations of monoamine

Figure 1: Prozac is a monoamine antidepressant known as a selective serotonin reuptake inhibitor (SSRI). SSRIs prolong the effect of serotonin in the synapse by preventing its “reuptake,” or reabsorption, into the neuron.
neurotransmitters cause depression (6).

Researchers often use antagonists, compounds that prevent molecules from binding their receptors and reduce molecules’ efficacy, to assess how compounds affect the body. Studies evaluating the effects of reserpine, a monoamine antagonist developed to treat high blood pressure, provide evidence to support the monoamine hypothesis (8). Research showed that, while effective in lowering blood pressure, reserpine causes MDD as a side effect and, as a result, is now rarely prescribed (8). Reserpine demonstrated that decreasing the efficacy of monoamine transmitters induces depression (8).

Based on this theory, newer medications, which have fewer or less severe side effects, were developed. These new treatments include antidepressants that selectively inhibit the reuptake of norepinephrine (NRIs), serotonin (SSRIs), and serotonin and norepinephrine (SNRIs) in the brain (7). Further support for the monoamine theory and the use of reuptake inhibitors as a treatment for MDD comes from clinical studies showing that patients with decreased concentrations of a serotonin metabolite were more likely to commit suicide (8).

However, there are a few aspects of depression that the monoamine hypothesis does not explain. Monoamine antidepressants require three to six weeks to reduce symptoms, despite producing stable neurotransmitter concentrations in the brain within a week of treatment (5). Additionally, they are only effective in treating about 50 to 60 percent of depressed patients (5). These discrepancies suggest that there are other factors aside from monoamine concentration that contribute to MDD onset and development.

The Glutamate Hypothesis

The glutamate hypothesis is one theory currently being researched that may help to supplement the monoamine hypothesis. Glutamate is an excitatory neurotransmitter that plays a key role in learning and memory (6). As early as 1982, reports have linked the overexpression of glutamate to depression (5). Since this finding, studies have shown that glutamate metabolism differs between people diagnosed with MDD and those without MDD, and that antagonists of various glutamate receptors produce antidepressant-like effects (5). However, approximately 30 percent of the brain expresses glutamate receptors, making it difficult to isolate the specific pathways through which glutamate contributes to MDD (6). Nevertheless, new research and treatments developed using the glutamate hypothesis show promise in alleviating the burden associated with MDD.

Gender Disparities

Although MDD affects people of all ages, ethnicities, and social and economic groups, it tends to be more prevalent among women, who are nearly twice as likely to develop MDD as compared to men (4). Women with MDD are also more likely to experience changes in their sleeping and eating habits; however, a larger proportion of depressed men tend to have suicidal thoughts under stress (4). The glutamate hypothesis provides a better explanation than the monoamine hypothesis for why symptoms and prevalence of MDD vary between genders.

A recent study from the University of Illinois at Chicago showed that glutamate receptor expression differs between genders, which may explain the gender disparities in MDD prevalence and symptoms (9). Researchers examined gene expression in MDD and control subjects without MDD postmortem. Initially, they found that those with MDD possessed more glutamate receptors in the brain, providing additional evidence for the glutamate hypothesis (5). Further data analysis revealed that this difference was mostly seen in female subjects (5). While women with MDD had significant increases in expression of eight glutamatergic genes as compared to controls, men did not (5). In fact, one glutamate receptor gene, GRM5, was seen in lower concentrations in men with...
“Although MDD affects people of all ages, ethnicities, and social and economic groups, it tends to be more prevalent among women, who are nearly twice as likely to develop MDD compared with men. Women with MDD are also more likely to experience changes in their sleeping and eating habits; however, a larger proportion of depressed men tend to have suicidal thoughts under stress.”

MDD as compared to male controls. While further research is necessary to determine the significance of varying glutamatergic gene regulation, the data provide a starting point in explaining why depression affects men and women differently (9).

**Mechanisms**

Until recently, the precise mechanism by which drugs that target glutamate receptors work was unknown. A recent paper published in *Molecular Psychiatry* provides valuable insight into glutamate receptors’ role in producing MDD. Previous studies showed that the protein p11 increases the sensitivity of serotonin receptors and helps generate the same behavioral effects seen after treatment with SSRIs (14). Low concentrations of p11 were linked to depression, and researchers increased the efficacy of SSRIs in MDD treatment by adjusting p11 concentrations (14).

New data reveals that p11 also regulates expression of the glutamate receptor mGluR5, linking the monoamine and glutamate hypotheses. Knockout mice lacking mGluR5 or p11 in glutamate neurons exhibited a lack of motivation to retrieve food pellets, a standard measure of depression and anxiety in mice models (15). The removal of mGluR5 or p11 from neurons sensitive to GABA, an inhibitory neurotransmitter, reduces these depression-like symptoms (15). These seemingly contradictory results can be explained by a complex regulatory effect produced by p11 and mGluR5; researchers found that an mGluR5 antagonist caused a decrease in excitatory (glutamate) neuron activity and an increase in inhibitory (GABA) neuron activity, mimicking the antidepressant effects of p11 (15). These results support previous studies linking depression to an imbalance between excitatory and inhibitory signals in the brain and provide evidence that mGluR5 plays a role in regulating neurotransmitters involved in major depression (15). This new information may help in the development of more effective treatments.

**Ketamine**

One drug currently being developed using the glutamate hypothesis is ketamine, also known as the party drug “Special K.” Ketamine has been shown to be a fast-acting antidepressant that can relieve depressive symptoms within two hours and continue to act for up to two weeks, even at low doses, making it a potentially useful drug for suicidal patients (9). However, ketamine is highly addictive and often abused as a recreational drug due to its ability to produce intense hallucinations and delusions (10).

Ketamine is an ionotropic glutamatergic N-methyl-D-aspartate receptor (NMDAR) antagonist. In other words, ketamine prevents glutamate from binding to its receptor on certain neurons, blocking its effect (9). Because many areas of the brain express glutamate, ketamine’s acute effects are difficult to study and control. In an attempt to reveal ketamine’s mechanism of action, a team of researchers at the University of Texas, San Antonio has isolated a brain circuit involved in producing ketamine’s antidepressant response without any of its negative side effects (10). This pathway contains the hippocampus and prefrontal cortex, which have both been shown to be involved in MDD (10). Their finding presents a new target for the development of fast and effective antidepressants with relatively few side effects.

The team used lidocaine, a nerve blocker, to inactivate the ventral hippocampus-medial prefrontal cortex (vHipp-mPFC) pathway in mice before administering ketamine (11). They...
then subjected the mice to the forced swim test, a standard procedure for measuring depression in mice based on their motivation to swim when placed in water (11, 12). Results indicated that inactivation of the vHipp-mPFC pathway prevented subjects from experiencing the long-term addictive effects of ketamine but not its acute antidepressant effects. Furthermore, activation of the pathway using designer drugs induced an antidepressant response similar to that induced by ketamine, while activation of other pathways did not. According to the lead author, Flavia R. Carreno, Ph.D., this pathway contributes only "to the beneficial effects of ketamine," while "another part contributes to its abuse and effects such as hallucinations" (10). The next step in developing a treatment based on this discovery is to identify a drug that selectively activates this circuit to harness its antidepressant effects (10).

Conclusion

The original theory of MDD states that it develops as a result of decreased concentrations of monoamine neurotransmitters in the brain. However, the monoamine hypothesis has proven insufficient in explaining all aspects of depression, causing scientists to develop additional theories to supplement the monoamine hypothesis.

Much of this research is directed toward the excitatory amino acid glutamate; high concentrations of glutamate correlate to MDD onset. Research suggests that the glutamate hypothesis may better explain how gender affects both the prevalence and symptoms of MDD, as well as provide a novel target for treating depression with fast-acting drugs such as ketamine. While depression still presents a large social and economic problem to our society, these new studies can help relieve this burden and reduce depressive symptoms in millions.  

References


CONTACT BREnda MIAO AT BRENDA.Y.MIAO.19@DARTMOUTH.EDU
The Birth of Fractals

Most people have probably seen examples of fractals, highly complex images in which the same geometric forms are repeated. They are often used as computer screensavers or in digital art, but there is more to fractals than their visual appeal. In the late 20th century, Benoît Mandelbrot coined the term fractal to describe mathematical and natural objects that had previously been considered oddities or paradoxes (1). Mandelbrot then formalized the study of fractals and discovered some of their deeper mathematical properties (1).

Examples of Fractals

One of the largest motivating forces behind fractals as a mathematical concept was the natural world. Mandelbrot’s book, *The Fractal Geometry of Nature*, held many examples of fractals in the natural world. In it he writes, “Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line” (1). He was particularly interested in coastlines and used them early in the book to define key concepts such as fractal dimension. Mandelbrot realized that jagged coastlines do not behave like regular lines, as they increase in length depending on how closely their edges are approximated. This led him to conclude that coastlines are not quite one-dimensional in a sense (1).

Although nature inspired Mandelbrot’s study of fractals, true fractals do not occur naturally, but are instead purely theoretical objects. The Koch snowflake (Fig. 1) is a snowflake shape constructed by starting with an equilateral triangle, then replacing the middle third of each side with a scaled down version of the original triangle. The process is then repeated on the resulting 12-sided shape so that each of the sides has their middle third replaced with a triangle that is one-ninth of the original size. This replacement rule is theoretically implemented an infinite number of times, making the Koch snowflake a purely mathematical construct. The Cantor set is another common example of a fractal. It is constructed by starting with a line and removing its middle third, then removing the middle third of the two remaining line segments and repeating this process (2).

Defining Fractals

Fractals all share some key properties, the first of which is self-similarity at different scales. In other words, fractals contain smaller copies of themselves. Furthermore, the self-similarity is never ending; the copies will continue to emerge no matter how far one zooms (3).

The second shared property has to do with their complexity and tendency to fill space. This property can be used to produce a precise definition of fractals, but some other definitions are required in order to fully understand them. Mandelbrot’s definition of fractals comes from two different descriptions of spatial dimension: the topological and Hausdorff dimensions, which will be defined...
in the following paragraph (1). He defines fractals as the spaces for which those two notions of dimension are not equivalent (1).

Dimensionality is a concept that is much easier to understand intuitively than formally. A space’s topological dimension is related to how it can be covered. Roughly speaking, a space is said to be n-dimensional if it can be covered by small sets so that each set overlaps a little bit with its neighbors, but no n+2 of the covering sets intersect. This concept is more easily explained visually. In Fig. 2, there is a line covered by circles and a plane covered by rectangles. Note that no three circles and no four rectangles intersect at any point, but there is no way to rearrange the squares so that no three of them overlap. Therefore, the line is 3-2=1 dimensional, and the plane is 4-2=2 dimensional (2).

An intuitive understanding of dimensionality can break down when dealing with fractals. For example, the Koch snowflake is a line, albeit a very convoluted one. Therefore, in one sense, it is one-dimensional. However, due to its fractal properties there is an infinite distance between any two points on the curve (2). Discrepancies such as these caused Mandelbrot to consider the Hausdorff dimension to further his study of fractals (1).

For a simplified understanding of the Hausdorff dimension, imagine a cube with side length s. Such a cube can be filled in by \( k^d \) smaller cubes of side length s/k. Similarly, a square could be covered by \( k^d \) squares that have been scaled in the same way. The dimension, \( d \), of each shape is the scaling factor for \( N \), the number of small copies needed to fill the original. In mathematical terms, \( N=k^d \) or \( d=(\ln N)/(\ln k) \). This idea can be extended to fractals with very interesting results. Consider the top section of the Koch snowflake (Fig. 3); in the bottom panel of the figure, four copies of the bolded section are clearly required to cover the original. However, the bolded section is only scaled down by a factor of three, meaning \( d=(\ln 4)/(\ln 3) \approx 1.26 \). This value for \( d \) is quite strange, as most people assume that dimensions must be whole numbers (4). However, these non-integer dimensions behave in a way one might expect in relation to whole number dimensions. For example, the Koch snowflake is something in between a line and a plane: because it is so jagged it takes up more “room” than a line, but it still does not “fill in” two-dimensional space like a plane would (1, 2).

The Koch snowflake has an infinitely long perimeter, as there are an infinite number of triangles along its border. The Koch snowflake must have a finite area, however, since it is bounded. For example, the area of the paper on which it is printed limits the Koch snowflake. Similarly, in a certain sense, the Cantor set has a length of zero because the length of each surviving line segment approaches zero. This means that the length of the removed parts equals the length of the original line, but somehow the set still contains points (for example, one-third will never be removed). Even more surprising is that, at any given point in the Cantor set, there are actually infinitely many points also in the set that lie within an arbitrarily small interval centered on the first point (5).

The Mandelbrot Set

The Mandelbrot set is probably the most widely known example of a fractal.
and is the reason most people know the name Mandelbrot. The set is defined using complex numbers, objects of the form $a+bi$, where $a$ and $b$ are regular, real numbers, and $i$ is defined as $\sqrt{-1}$. Complex numbers can be represented as points on a plane (Fig. 4), where the $x$-axis represents the real component of a complex number, and the $y$-axis represents the imaginary component. The magnitude of a complex number is defined as the distance from the point on the complex plane to the origin and can be calculated using the Pythagorean theorem (6).

The Mandelbrot set is defined using an iterated function, which hints at its fractal nature. The function itself is straightforward: $f(z)=z^2+c$, where $z$ and $c$ are complex numbers. The function is then iterated, meaning that each output is used as the next input, with the first input being zero. The resulting series of outputs can be labeled $z_0, z_1, z_2, \ldots$ and they follow the rule: $z_{n+1}=f(z_n)$ and so on. Given that the value for $z_0$ is zero, there are two possibilities for the series of outputs that depend solely on the choice of $c$: it can either go to infinity or it can remain bounded. The Mandelbrot set is defined as the set of values of $c$ for which the sequence of outputs remains bounded under the iteration of $f(z)=z^2+c$. For example, if $c=1$, each output will be an integer that is strictly larger than its predecessor, creating a series that will grow without bound. However, if we chose $c=-1$, then the series will alternate between negative one and zero and remain bounded, so negative one is in the Mandelbrot set, and one is not (7).

It may seem impossible to determine if a series is actually unbounded or if it simply has a very high upper bound. However, if the magnitude of any term in the series is greater than two, the series will always go to infinity (8). Computer programs can use this fact to generate images of the set, as they can disregard any points of which the series has a value outside of that bound. Typically, computer-generated images will have regions of different colors around the set (Fig. 3). The color at each point refers to the number of iterations needed for that value of $c$ to generate an output with magnitude greater than two (9). The Mandelbrot set is connected, which means that one can draw some path between any two points in the set so that the path is always inside the set; technically this is a definition of path-connectedness, a slightly stronger condition. Mandelbrot believed that the set did not have this property because computers in his time could not generate detailed images of the set, leading him not to notice the thin filaments that connect its larger sections (10).

### Conclusion

Unusual or interesting topics lead mathematicians to create “new” math in order to understand them. After this “new” math is created, it often takes on a life of its own and can yield countless new questions and applications. Fractal geometry is now more than just a mathematical curiosity: it has applications in seemingly unrelated fields such as finance, fluid dynamics, and image processing, with yet more applications waiting to be discovered (11).

**CONTACT KRISHAN CANZIUS AT KRISHAN.CANZIUS.18@DARTMOUTH.EDU**

**References**

Deep Learning: Can Computers Learn Like We Do?

BY CARA VAN UDEN ’19

Computational Chess Masters

Almost 20 years ago, IBM’s Deep Blue supercomputer beat the reigning world chess champion, Gary Kasparov, for the first time (1). Since then, chess-playing computers—also called chess engines—have become significantly stronger, leaving the best humans little chance against even a smartphone. While computers have become faster, the way chess engines work remains unchanged. Their power relies solely on searching through all possible future moves to find the best one.

No human can match that level of brute force. While Deep Blue searched 200 million positions per second, Kasparov probably searched no more than five per second, yet he played at essentially the same level (1). Clearly, humans have a trick up their sleeve that computers have yet to master. Unlike computers, people can evaluate chess positions to narrow down the most profitable avenues of search. This technique dramatically simplifies the computational task, as it prunes the tree of all possible moves to one of just a few branches.

Computers have never been good at this type of “prescient” computing, but they are steadily improving due to the work of Matthew Lai at Imperial College London. Lai has created an artificial intelligence machine called Giraffe that has taught itself to play chess by evaluating positions in a way similar to humans (2). After only 72 hours of training, the new machine plays at the same level as the best conventional chess engines, many of which have been fine-tuned by chess grandmasters over many years (2). On a human level, it is equivalent to FIDE International Master status, which is within the top 2.2 percent of tournament chess players (2).

The technology behind Lai’s new machine is a neural network modified through a technique called deep learning. Inspired by the human brain, it consists of several layers of processing nodes, called neurons, that are connected in ways that change as the system is trained. This training process uses massive data sets—Lai’s research used 175 million possible chess positions—to fine-tune the connections so the network produces a specific output given a certain input (2).

Building a Brain

Traditional methods of computer programming require programmers to feed computers with information and rules about the world through tedious lines of code. This approach limits systems to controlled but narrow applications, such as phone menu systems that ask one to make queries by saying specific words (3).

Lorenzo Torresani, a computer science professor at Dartmouth, explained that deep learning turns the traditional approach on its head by expanding upon machine learning, which describes any approach that allows computers to learn to solve a problem on their own by seeing examples of solutions (4).
Normally, programmers have to tell a computer the exact steps to solve a problem. With machine learning, Professor Torresani explained, computers learn the steps for themselves (4). Until recently, machine learning could only tackle a small range of tasks, particularly those that use structured data, that is, data in rows and columns, such as what might come from a database table. However, most data that humans work with every day is not structured. Instead, it is in unstructured forms like images, sounds, and natural language (4).

Deep learning is a technique that brings machine learning to unstructured data; it is inspired by a very simplified form of how the brain works (4). Deep-learning software attempts to mimic the activity in layers of neurons in the neocortex, the wrinkly 80 percent of the brain where thinking occurs (4). The software learns to recognize patterns in digital representations of sounds, images, and other data (4).

In a deep neural network, a program maps out a set of virtual neurons in several layers and then assigns random numerical values, or “weights,” to connections between them (4). The weights determine how each simulated neuron responds to a digitized feature, such as an edge or a particular frequency in a sound (4).

Professor Torresani also explained that the first layer of neurons learns primitive features, like the aforementioned edge in an image, through finding combinations of digitized pixels or sound waves that occur more often than they should by chance (4). Once the first layer accurately recognizes those features, they are fed to the next layer. This new layer trains itself to recognize more complex features, like the subject in an image or a combination of speech sounds (4). The computer repeats the process in successive layers until the system can reliably recognize phonemes, which are the individual unit of sound in spoken syllables, or objects (5).

Some of today’s artificial neural networks can train themselves to recognize complex patterns. Programmers train a neural network to detect an object or phoneme by exposing the network to images containing objects or sound waves containing those phonemes (3). A special algorithm, called a back-propagation algorithm, adjusts the weights until the network consistently recognizes a particular pattern of pixels or phonemes (3). Deep neural networks can accurately recognize faces, process language, and potentially diagnose diseases.

The basic idea that software can simulate the neocortex’s large array of neurons in an artificial “neural network” is decades old, and it has led to as many disappointments as it has breakthroughs. However, improvements in mathematical formulas and increasingly powerful computers have enabled computer scientists to model more layers of virtual neurons than ever before (3). Deep learning is now becoming state of the art in areas such as image processing and drug discovery.

**From Cats to Computer Vision**

In 2011, Google demonstrated Google Brain. One of the largest neural networks yet, the system currently has more than one billion connections. A team led by Google Fellow Jeff Dean and Stanford computer science professor Andrew Ng showed the system images from 14 million randomly selected YouTube videos (6, 7). One simulated neuron in the software model fixated on images of cats. Other neurons focused on human faces, yellow flowers, and other objects. With the power of deep neural networks, the system identified these discrete objects even though no humans had not defined or labeled them (6, 7).

Google Brain excels in image recognition. Though it only correctly categorizes objects 16 percent of the time, it is still 70 percent better than previous methods (8). In addition, there are 22,000 categories for the Google Brain to choose from (8). Correctly slotting objects into some of these categories requires such things as distinguishing between two similar varieties of jellyfish, a task that would be challenging even for most humans. When the system was instead asked to sort images into 1,000 broader categories, the accuracy rate jumped above 50 percent (8). Google researchers believe the Google Brain has broad applications across technology— it will help with the company’s search algorithms, self-driving cars, and boost Google Translate, among other functions.

Microsoft is also exploring deep learning and recently announced advances in its technology after winning a computer vision competition in 2015.

Figure 2: Deep learning could eventually be used to diagnose diseases by analyzing MRI images such as the one below.
In December 2015, Microsoft won first place in several categories in the ImageNet and Microsoft Common Objects in Context Challenges (9). The competition looked at approaches to recognizing images, like objects, faces, and emotions, in both photographs and videos. Jian Sun, the principal research manager at Microsoft Research who led the project, said that the company’s system outperformed its competitors and won all three categories, classification, detection, and localization, with a single type of neural network, defeating its competitors by a large margin (9).

According to Microsoft, the researchers trained the deep neural network system with eight layers three years ago. Last year Microsoft used 20 to 30 layers, but their newest system has 152 layers—five times deeper than any previous system (9).

Neural networks are built in a series of layers. Theoretically, more layers should lead to better results, but, in practice, back-propagation signals vanish as they pass through each layer, eventually leading to difficulties in training the whole system. Microsoft’s researchers discovered a way to add more layers and still obtain accurate results, eventually developing a system called “deep residual networks” (9). The deep residual net system they used for the ImageNet contest uses this new “residual learning” principle to guide network architecture designs (9). Residual learning alters the learning procedure by redirecting information flow, allowing information to skip through multiple neural layers at a time (9).

Returning to Dartmouth, Professor Torresani works to extend a deep learning approach to video processing (4). Video, composed of thousands of still images, is an even more massive set of data, but the extractable information from videos has wide-ranging applications, from motion capture to modeling human movement. Professor Torresani currently works in motion capture to reconstruct the continually changing shape of a non-rigid object in three-dimensions from its two-dimensional motion as seen in a video (4). Using motion capture data, his team has developed methods to model the stylistic differences with which people perform specific actions, like running or walking (4).

**Siri, Speech-to-Text, and Translation**

The current goal of natural language processing research is to help computers understand and even speak in natural language. Researchers are trying to provide computers with enough understanding of natural language to do useful things, like search the internet or answer questions.

Already, deep learning has improved voice search on smartphones. Until last year, Google’s Android software used a method that misunderstood many words. In preparation for a new release of Android last July, Google researchers replaced part of the speech system with Google Brain. As a result of the multiple layers of neurons that allow for more precise training on the many variants of a sound, the
system could recognize scraps of sound more reliably, especially in noisy environments such as subway platforms (7). Almost overnight, the number of errors fell by up to 25 percent (7). The results were so notable that many reviewers now deem Android’s voice search superior to Apple’s more famous Siri voice assistant (7). In addition, Microsoft has deployed deep learning in its Windows Phone and Bing voice search (5).

In the written world, natural language processing enables computers to conduct live analysis of plain text, such as emails, Word documents, and slideshow presentations. This ability has the potential to lead to features like automatic fact checking, computerized citations, and footnotes.

Deep-learning models can use phoneme data from English to train systems to recognize the spoken sounds in other languages more quickly (6). This capability has already created dramatic improvements in translation services. People who do not share a common language may soon be able to have live, online conversations without the use of a translator. In 2012, Microsoft CRO Rick Rashid impressed attendees at a lecture in China with a speech software demonstration that transcribed his spoken words into English text with an error rate of seven percent, translated them into Chinese-language text, and then simulated his own voice speaking them in Mandarin (6).

Diagnosis and Drug Discovery

In 2012, the pharmaceutical company Merck offered a prize to whomever could beat its best programs for predicting useful drug candidates. Researchers from the University of Ontario and the University of Washington used a deep learning system to search through database entries on more than 30,000 small molecules, each of which had thousands of chemical property descriptors (10). The researchers then used the system to predict how each small molecule acted on 15 different target molecules and find an effective drug agent (10). Though the team designed its software with no specific knowledge of how molecules bind to their targets, they still won $22,000 and improved on Merck’s baseline by about 15 percent (6, 10).

On the clinical side of medicine, the startup Enlitic is developing deep learning methods to scan and process images, such as those produced by CT or MRI scans (11). Researchers believe that, by examining these images for diseases, brain tumors, and cancer, among other ailments, physicians will be able to detect these illnesses more quickly and accurately. Enlitic will gather data about a particular patient, such as medical images, lab test results, and doctors’ notes, and its deep learning algorithms will analyze the data to reach a diagnosis and propose treatments. This approach may uncover previously unnoticed patterns within medical data and discover the underlying mechanisms of many diseases (11).

Looking to the Future

Deep learning has the potential to revolutionize many areas of technology, from medicine to computer vision. Training deep neural networks requires massive amounts of data, however, which can be difficult to obtain. Privacy concerns can arise when companies use individuals’ data, such as photos and medical records, as part of a training data set. In the 21st century, information has become a commodity, and private data is often less private than previously thought. However, data has also become more openly accessible, allowing researchers and companies to introduce deep learning to more applications than ever before. In the future, it will be important to balance deep learning research with awareness of online privacy.

CONTACT CARA VAN UDEN AT CARA.E.VAN.UDEN.19@DARTMOUTH.EDU

References

Deep Sequencing: The Rapid Genomics Revolution

BY SAM REED ‘19

Deep Sequencing

The genome, the whole of an organism’s DNA or RNA, is comprised of four distinct nucleotides, or bases, which recur many times in a sequence. In the human genome, this sequence is about three billion base pairs long, but that chain of bases, containing only those four nucleotide variants, determines whether an organism will be a human, a bacterium, or something else entirely.

Because the genome helps determine the phenotype, or observed characteristics, of an organism, it provides a tremendous amount of information (1, 2). Discoveries in genomics range from the origin of a human feature or disorder to viral characteristics that could be key in fighting an epidemic (1, 2). Genetic sequencing, the process of determining the sequence of bases in a genome, enables these discoveries (3). However, for a number of reasons, the sequencing of a small random stretch of a genome is neither practical nor ideal, making sequencing on a larger scale necessary.

Deep sequencing, the sequencing of an organism’s entire genome, is a potential solution to this problem and provides information that is useful to many different fields (1, 4). It was not until this past decade, however, that the technology needed to make this method viable for clinical use even existed (5).

While smaller genomes were sequenced in the late 20th century, the human genome was not fully sequenced until the Human Genome Project began doing so in 1984 (3, 5). The Human Genome Project published the first drafts in 2001 and completed the project in 2003 (3, 5). The project consumed 13 years, three billion dollars, and the work of scientists across the world. It was also one of the first projects to use the Sanger method of nucleotide sequencing (3).

The Core Technology

Sanger sequencing was the original form of deep sequencing and is still used today to examine shorter sequences, such as single genes (5). In Sanger sequencing, single stranded DNA is mixed with primers, DNA polymerase, and normal nucleotides, all of which are needed to form a complementary strand (5). Modified nucleotides lacking the 3’ hydroxyl necessary to form a bond with the following nucleotide are also added; these modified nucleotides cause DNA synthesis to terminate, creating DNA strands of varying length. After repeating this process many times, a collection of DNA fragments can reveal the nucleotide at each position, allowing the overall sequence to be determined (5).

Initially, before the invention of fluorescent bases, Sanger sequencing involved a different reaction for each of the four nucleotides so that chain termination could be associated with the presence of a specific nucleotide. But Sanger sequencing, before and after fluorescent bases, had several inherent problems: 1) the fragments were as short as a few bases long,
and 2) the technology of the time could only sequence as many as a few hundred fragments at once (5). Sanger sequencing, as a result, can only accurately sequence as many as 1000 bases. With such a method, a three billion-base pair genome requires huge amounts of time and massive DNA libraries (5). The time and money the Human Genome Project required is a testament to the difficulties of large scale Sanger sequencing.

Until recently, the clinical use of deep sequencing was neither practical nor possible. At the time of the Human Genome Project, three billion dollars would have been insufficient to sequence the genome of a patient in a timely manner, leaving the final product useless once obtained. For deep sequencing to be used in a clinical setting, the technology would not only need to be widely available, but also fast and affordable (6, 7, 8). At this stage, next generation sequencing (NGS) technology entered the market.

NGS technology performs deep sequencing in a high throughput and massively parallel manner, meaning that many fragments are sequenced in unison to generate large amounts of data (3). This technology allows what once took hundreds of people and hundreds of machines to be performed by one person on one computer in less than a day (3, 5). Since the first NGS-level sequencer was released by Lynx in 2000, NGS has advanced exponentially into the viable tool that it is today (5, 8).

Next-generation sequencing begins with the random fragmentation of DNA. Next, nucleotide sequences called adaptors are bound to the ends, creating the “library” (9). Afterwards, the fragments are moved into a flow cell, where promoters complementary to the adaptor sequences are present in a single layer (9). The fragments bind to the promoters as single strands (9).

Each fragment is then cloned through bridge amplification, whereby the fragment bends into a “bridge” as a result of adaptors binding promoters in the flow cell, and DNA polymerase and nucleotides in the surrounding solution create the complementary strand (9). The fragments are then denatured, separating the strands, and the process is repeated until each fragment is surrounded by a “cluster” of its clones (9). Finally, the fragments are sequenced by adding fluorescently labeled bases into the cell that bind the single stranded fragments (9). Digital images are taken of each “cluster,” and the order of bases is determined from the light emissions recorded (9). Finally, bioinformatics software realigns the fragments and analyzes the resulting sequences (9).

**NGS: Advances and Impacts**

Within the last decade, there have been a number of advances in NGS technology. Multiplexing, for example, allows different strands of DNA to be sequenced simultaneously by attaching different distinctive adaptors to the fragments (9). One of the most significant advances within the last decade is the steady increase in the throughput of computer technology (6). These advances and others progressively increased the speed of subsequent NGS models; what was once possible in 13 years can now be accomplished in a matter of days (10).

In January 2015, Illumina released the HiSeq X, which is in many ways exemplary of the advances in deep sequencing technology that occurred in the decade prior to its release (11). The HiSeq X can sequence 20,000 human genomes within a year if run at full capacity (11). Additionally, the HiSeq X decreased the price per genome sequenced significantly, costing only $1,000 per genome sequenced, nearly a tenth of prior costs (1, 9, 11).

However, in terms of its low cost, the HiSeq X is not an exception to the rule. The cost of deep sequencing has decreased drastically in the last decade. Since 2007, the costs have
fallen far below the curve projected by the National Human Genome Research Institute, which predicted a price per genome of just under $1,000,000 by 2015 (12). Even before the HiSeq X, the price was under $10,000 (11).

NGS has spread to labs around the world, made possible by its increasing effectiveness and viable costs (1). Higher availability means higher overall use and impact. Today’s lower costs also make it possible to sequence the DNA of many individuals in a short time period, allowing for larger sample sizes and more flexibility in experimental design (4, 5).

Advances in deep sequencing have also impacted medicine. By the year 2011, 3,000 single-gene Mendelian genetic disorders had been identified, which would have been entirely impossible without NGS technology (8). As a result, children with rare genetic disorders who would have gone undiagnosed in the past receive a diagnosis and, possibly, the appropriate treatment (11). Deep sequencing has also been helpful in treating cancer. Deep sequencing enables doctors to compare the current genome of a patient to his or her original genome, revealing areas where mutations may have occurred. The rapid speed at which deep sequencing can occur, along with its affordability, also allows for multiple patients to receive such treatments (7).

Deep sequencing has revolutionized the way we treat certain diseases. Two companies, Foundation Medicine and Genomic Health, have been integral to this change. Foundation Medicine, for instance, sequences the tumors of patients with colon cancer and then prescribes a treatment that has proven effective against other tumors with those specific mutations (11). NGS has also been immensely helpful in tracking and treating infectious diseases. Treating viral diseases like hepatitis C is possible with Sanger sequencing, but NGS is needed for pathogens with larger genomes (13).

NGS has also been a major player in fighting epidemics like the recent Ebola outbreak (2). Because viruses mutate quickly, two patients who are found to have the same viral genome most likely contracted it at a similar time and location (2). NGS allows researchers to sequence the viral genome and go back to the disease’s source to learn who else might have the disease and who is in danger of contracting it (2).

NGS is also a powerful tool for curing and preventing disease. Because vaccines typically contain a milder version of the virus or viral proteins made in a lab, sequencing the viral genome helps to inform vaccine development and prevent the spread of disease (2). This same sequenced information allows researchers to determine the form and function of the virus’s key proteins, which is helpful in antiviral drug development (2).

An Incomplete Revolution

While deep sequencing has revolutionized genomics and enabled previously unimaginable scientific feats, there are a number of factors that have impeded the technology’s proliferation and overall impact. With regards to the technology itself, the short “read length,” the length of fragments the machine can read, is a threat to its accuracy. While NGS allows for high-throughput and parallel sequencing, the read length is actually shorter than that of Sanger

*Image courtesy of NIAID. Available at https://www.flickr.com/photos/niaid/16441626349*

*Figure 3: NGS played a significant role in fighting the recent Ebola epidemic.*
sequencing (7). Higher read length means less gaps in the genome and a better digital assembly of the entire genome (6).

While the cost to sequence a genome is far lower than it has ever been and is still decreasing, the high cost of sequencing machinery prevents some institutions from seeing the best of NGS. The HiSeq X, for example, costs one million dollars and must be purchased in sets of 10, limiting its use to high-volume research institutions with ample funding (11). While the the NextSeq 500 is priced for organizations such as hospitals, it can only process one genome at a time (11).

The need for advances in bioinformatics, however, has been referred to as the largest bottleneck in NGS implementation (8). Not only does bioinformatics require a relatively large monetary investment, but it also requires training in interpreting large molecular data sets. Among physicians, there are only 1,000 or so medical geneticists and 3,000 genetic counselors with this level of training in the US (12). While these numbers are not insignificant, they are still not high enough to meet the demand for genomics testing in the US (12).

Solutions to these problems are possible, however, and, in some cases, right around the corner. Over the last decade, the read lengths in NGS technologies have steadily increased (6). In addition, there are 17,000 pathologists in the US alone who could be trained sufficiently in bioinformatics (12).

Nanopore technology, which was only recently made available, is bringing new levels of accessibility and usability to deep sequencing. In this new kind of sequencing, a protein channel, or “nanopore,” lies on a membrane (14). An electrical potential is then passed along the membrane, and the current experiences different disruptions based on the molecule that passes through it (14). A strand of DNA can be sequenced by passing it through the nanopore and monitoring the fluctuations in potential that its bases generate (14). Sequencing can occur without library formation or amplification.

A notable result of nanopore sequencing technology is the MinION, produced by Oxford Nanopore Technologies. The MinION is a palm sized nanopore deep sequencing machine that can be taken into the field and can sequence a genome in under 48 hours (15). Researchers have tested this device extensively in West Africa, and NASA scientists hope to take it into space (15). While the MinION still has a number of unknowns, the machine’s price is predicted to be very low, making it even more accessible (15).

One Giant Leap for Mankind

Just over a decade ago, sequencing one human genome took 13 years and three billion dollars. Today, scientists can sequence 20,000 genomes in one year by using machine for $1,000 per genome. Handheld machines that can be used in the field are coming onto the market. The technology that revolutionized the study of the human genome is becoming increasingly accessible, broadening the frontier for new discoveries and health work in genomics.

References


“The need for advances in bioinformatics, however, has been referred to as the largest bottleneck in NGS implementation.”
A Smooth Landing

A pigeon bobs its head when it lands. This is a simple observation, but the explanation for the pigeon’s behavior is actually quite surprising. The pigeon bobs its head for the exact same reason that a squirrel runs back and forth along a fence: to better judge distance. When humans move their heads back and forth, foreground objects appear to move more than background objects. This effect is called motion parallax, and it is the only reason pigeons can land without crashing (1).

In contrast to pigeons, humans rely mostly on binocular vision, or stereopsis, to judge depth (1). By comparing two images of the same scene, each taken from a slightly different perspective, one can triangulate distances with great accuracy. Although this binocular vision gives an excellent estimate of depth, it decreases an organism’s field of view because the images from each eye overlap. For squirrels and pigeons, a wide field of view is of paramount importance to detect potential predators. Humans and other tertiary predators use binocular depth perception because it gives them an advantage when hunting. Each organism measures depth in a manner that best suits its particular needs.

Tools of the Trade

A broad array of visual cues exists for measuring depth of field. Textures appear blurrier when viewed from a distance. An object of known size shrinks predictably with distance, and closer objects occlude more distant objects. Very distant objects will appear more blue as a result of light’s interaction with the atmosphere.

Some painters, most notably Cézanne, took advantage of this ultraviolet shift to paint with more depth.

How does the brain use each of these effects to measure depth? Recent neuroimaging studies point to the occipital lobe, which is where humans process visual stimuli. Raw visual information first enters the lobe by traveling from the eyes and along the optic nerve. Processing begins in the rearmost part of the lobe in an area called V1, or the primary visual cortex (1, 2). The neurons in V1 are organized spatially, and they process simple features such as dots and lines (1, 2). During this early stage of processing, signals from the two eyes begin to interact, which is crucial for binocular vision (2, 3).

Just in front of V1 lies V2, the secondary visual cortex. The secondary visual cortex recognizes slightly more complex features (2). For example, neurons in the V2 begin to discriminate between the foreground and the background of an image. Beyond V2, visual information begins flowing in two different directions. One is called the dorsal stream and the other is called the ventral stream (2). The dorsal stream helps the brain coordinate motion (2). Neurons in this stream have poor spatial resolution but process signals quickly and specialize in understanding motion. Depth perception cues derived from moving objects are processed in this stream (2). A batter judging the distance of an approaching baseball, for example, would use his dorsal stream. The ventral stream is much slower to process signals but helps the brain recognize complex features such as faces (2). Neurons in this stream specialize in recognizing color.
shapes, and edges, meaning that static depth cues such as gradients and converging lines are processed in the ventral stream (2). As a result of such processing, viewers can perceive depth in a photograph or painting.

Frontline Research

The visual system is far more complex than this working model of depth perception might suggest. In fact, neuroscientists do not yet understand some of its most fundamental properties. How does the brain use memories to perceive depth? Where in the occipital lobe does the brain process binocular vision?

Both memory and reasoning have been shown to play a role in depth perception. Familiarity with the size of a particular object, such as a chair, can allow individuals to infer the size of a new object placed beside the chair. Humans’ brains must be able to use their memory of chairs objects and the fact that the two objects have been placed beside each other to make inferences about the new object’s size.

If this is true, connections between the regions of the brain involved in reasoning and memory must exist, such as the prefrontal cortex and hippocampus, and areas of the visual cortex. Indeed, studies have found strong evidence for communication between the hippocampus and the ventral visual stream (4). However, understanding how the brain combines memory with visual processing is an area of ongoing research.

Modeling binocular depth perception in the brain is another conundrum. Neurons that combine signals from both the right and left eyes are known as binocular cells. Scientists initially believed that the complex transformations for building a three-dimensional scene from two images must happen in the last, most abstract steps of visual processing (4). A clever visual illusion called the random dot stereogram changed this view. To build a random dot stereogram, create two copies of a rectangle filled with a random distribution of black and white dots. One rectangle will be for the left eye, and the other will be for the right. Next, copy a small square from the center of the right rectangle and shift it a few pixels to the right. When a human’s eyes are crossed so that the two rectangles appear superimposed over one another, the square region suddenly pops into three dimensions.

The random dot stereogram is interesting because random dots do not contain any higher-level patterns and are perceived as “noise” by regions beyond the primary visual cortex (9). Binocular cells must operate at the level of V1 for this effect to occur (5). He interpreted these paradoxical findings to mean that binocular cells also operate at the level of V2 and above (5). Ramachandran explains the paradox with a “20-questions game” analogy: at each stage of visual processing, the brain sends signals back to the previous stage. Then, after successive iterations of this recursive behavior, the brain arrives at a solution (5).

One recent study of connectivity in the mouse visual cortex revealed that, in most areas, the number of feedback neurons balances with the number of feedforward neurons (6). This finding supports Ramachandran’s theory, as the large number of feedback connections might send information from later stages of vision processing back to the primary visual cortex.

Limits

The human eye is well known for pushing the limits of vision. Certain rods, for example, are capable of detecting just a single photon (7). Results are no less impressive for binocular vision. In fact, the visual system produces binocular vision that is sensitive to differences between the two eyes smaller than the width of a single cone photoreceptor (7). Such precision can only be achieved when the brain maps every feature in the left eye’s input to every feature not processed in V1 (5). Ramachandran was also able to build stereograms solely from higher-level patterns.
in that of the right eye. In terms of distance, binocular vision is most effective at distances of two meters or less, but it can function at distances of up to 40 meters (8).

There are two clinical tests for evaluating an individual’s binocular vision. The first is the random dot stereogram, discussed above. This test is useful because the random dots eliminate all other depth cues so the patient can only describe shapes in the image if he or she has binocular vision (9). The second test is called a contour stereotest. The most common test of this type is the Titmus Fly Stereotest, which unsurprisingly consists of a large image of a fly (9). When viewed with 3D glasses, different parts of the fly project to different depths. Depending on how a patient describes different parts of the fly, a physician can deduce his or her aptitude for depth perception (9).

Closing Thoughts

Take a step back and look at depth perception from a new angle. What makes it such a difficult problem to solve? Why does it happen at so many different levels of visual processing in the brain? The problem is difficult for two reasons: 1) given a large amount of visual information, humans must reduce all inputs to a few basic measurements, such as the relative sizes of two objects, and 2) in order to create a three-dimensional representation of the world from two-dimensional images, difficult mathematical transformations must also be performed. The brain solves this problem by combining many visual cues for depth. As discussed, some of these cues, such as binocular effects, must be processed as early as V1, others, such as the convergence of lines in the distance, happen in V2, and still others require high-level inferences about past experiences stored outside of the occipital lobe, such as in the hippocampus. The sum of all of these individual visual effects, each occurring at a different stage of the visual hierarchy, gives organisms what is known as depth perception.

CONTACT SAM GREYDANUS AT SAMUEL_J.GREYDANUS.17@DARTMOUTH.EDU

References


Figure 4: The human field of view. Humans can only estimate depth in the centermost 30 degrees.
What is Systems Biology?

For the past several decades, drug discovery has focused on identifying drugs that activate or inhibit particular molecular targets. As a result, scientists have generated a large amount of "omics" data, information related to genome and protein structure, function, and sequence collected to catalog individual molecular components in biological systems (1). The biological system is complex, however, as a distinct biological function cannot be attributed to an individual protein or gene. Biological functionality emerges from interactions among several hundred genes and proteins within the cellular environment (1). In order to understand living cells, it is first necessary to understand these networks and their underlying nodes and connections.

By considering disease as a condition involving perturbed cellular network connections, systems biology aims to gain a better understanding of the molecular basis of human disease and drug mechanisms of action on a deeper level (1). Through a comparison of normal and diseased networks, researchers can identify critical nodes and modulate them to reconfigure the perturbed network structure back to its normal state. Systems biology attempts to manipulate specific biological mechanisms through drug treatments under various conditions.

Top-Down Approaches

Systems biology approaches can be divided into two categories: top-down and bottom-up approaches. Top-down approaches analyze omics data sets, including genomics, transcriptomics, proteomics, and metabolomics, to identify the networks and logic behind a given biological function (1). These omics data sets display the "big picture" of the biological system. From this data, researchers then try to identify the pathways and network components responsible for the system's behavior from this data, with the ultimate goal of understanding the biological network well enough to create targeted drug treatments (1, 2). Examples of top-down systems biology approaches include gene-set enrichment analysis and the Connectivity Map, both of which start with big data and end with specific cellular pathways.

Gene-Set Enrichment Analysis (GSEA)

In the past, gene expression analyses focused on finding a single gene that showed differences between the normal and diseased states (2). Although insightful with respect to the role of a particular gene, gene expression analyses fail to consider interconnected biological processes controlled by networks of genes such as transcriptional and metabolic pathways (2).

Gene-set enrichment analysis (GSEA) provides an overview of the processes that change when a system is perturbed, as in the case of disease or drug treatment. GSEA focuses on gene sets that share a common biological function, chromosomal location, or regulation network (2). The first step in GSEA is to sort gene expression data sets according to their correlation
with phenotypes that are characteristic of a given condition. Next, researchers correlate ranked lists of regulated genes with pre-defined gene categories to identify pathways and biological processes that are altered the most in data sets of interest (3).

Aravind Subramanian of the Broad Institute of Massachusetts Institute of Technology and Harvard and his team applied GSEA to compare independently derived Boston and Michigan lung cancer data sets and used these comparisons to identify genes that signify poor outcomes in lung cancer (2). While single-gene analysis provided no clear correlation among genes, GSEA identified eight gene sets in the Boston data and 11 gene sets in the Michigan data that significantly correlated with poor lung cancer outcomes (2). There was also a prominent overlap among the enriched, or correlated, gene sets in the two studies: roughly half of the gene sets in the two studies showed similarities, and a few non-identical gene sets were related to the same biological processes (2). GSEA was able to produce greater correlational consistency between genes and poor lung outcomes in the two data sets than single-gene analysis. In the future, these capabilities may allow engineers to develop optimized drugs that target appropriate nodes in biological pathways.

Connectivity Map

A complex disease is not seen as a failure of a single gene or target. Rather, it is viewed as a consequence of disease-perturbed networks characterized by a specific disease signature (1). A complex disease is not seen as a failure of a single gene or target. Rather, it is viewed as a consequence of disease-perturbed networks characterized by a specific disease signature (1). As a result, the objective of successful drug treatments is to return the disease signature back to normal.

Researchers developed the Connectivity Map database with this very goal in mind. The Connectivity Map is a top-down approach that aims to describe biological states in terms of gene expression signatures, which include mRNA levels, DNA methylation patterns, metabolite profiles, and protein expression. This data is then used to generate large databases of drugs and genes for given biological states (4). By comparing the corresponding gene-expression signatures from this database, the Connectivity Map plans to uncover similarities among signatures (4).

Steven Kunkel’s team at the University of Iowa developed a potential therapy for skeletal muscle atrophy using this Connectivity Map approach (5). Focusing on fasting-induced muscle atrophy, they analyzed mRNA from fasting humans and fasting mice. This analysis provided two sets of differentially expressed genes, which the team then used to search the Connectivity Map to locate small molecule inhibitor candidates for muscle atrophy (5). The scientists identified ursolic acid, a compound found in apple peals, as a possible treatment for muscle atrophy, as its signature was opposite to those of atrophy-inducing stresses (5). Experiments verified that ursolic acid
reduced muscle atrophy and stimulated muscle hypertrophy in mice (5).

The Connectivity Map has encouraging application results, as gene expression signatures can be used to identify drugs with common mechanisms of action, discover unknown mechanisms of action, and recognize potential therapeutics (4).

**Bottom-Up Approaches**

Whereas top-down systems biology approaches start from data sets representing entire biological systems, bottom-up approaches start from a single component of a system. Many individual components are then gathered into a pathway to simulate the behavior of an assembly (1). Researchers use different modeling approaches depending on the nature of the data and the level of understanding of the studied biological system. If the kinetic parameters, such as binding affinity and reaction rate, and the concentrations of interacting molecules are well established, mechanistic modeling is the appropriate approach. If such knowledge is insufficient, Boolean-logic based modeling, which assigns one of two values to all possibilities, can be used. Like the top-down systems biology approaches, the bottom-up approaches are used to develop better drug treatments for disease.

**Mechanistic Models**

Mechanistic models are mathematical models that seek to understand the dynamics of biological processes that govern cellular functions (1). These models apply Michaelis-Menten kinetics and the law of mass action to define the rates of production and consumption of molecular components (1). Because a multitude of signaling pathways govern and coordinate the behavior of cells and many disease processes arise from defects in these signaling pathways, it is important to develop quantitative dynamic models that can identify effective drug targets.

Ordinary differential equations (ODEs) are most commonly used to model signaling pathways in biological systems as networks of biochemical reactions (1). The known kinetic parameters associated with each reaction must be fitted to experimental time-course data of specific cellular or physiological systems (6). ODE based models can quantitatively predict the time course for each component in the system over the simulation period, making them both appropriate and powerful analytical tools (6).

Lisl Shoda and her team at Entelos, Inc., a biosimulation company, applied mechanistic modeling approaches to represent type 1 diabetes pathogenesis in non-obese diabetic mice (7). Shoda and colleagues used ODEs to represent the pancreas and the pancreatic lymph nodes, as well as the interactions of multiple cell types involved in glycemic control. These ODEs allowed them to effectively simulate the outcomes of hypothetical therapeutic interventions in treating type 1 diabetes.

Molecular mechanistic models can identify and validate novel drug targets by simulating the expected effects of pharmacological interventions on cellular systems (8). A given drug’s effect may be described by adding a drug binding event to the model or changing the value of parameters that represent the drug target’s biological activity (8). It is also possible to design virtual drugs with completely new properties and test their effect on systems of interest (8). These results can help create the framework for a drug profile and enable drug discovery.

**Boolean-Logic Based Models**

Mechanistic models require information regarding all of the parameters used in the system, but such information is often unattainable. Logic-based modeling can be used in these cases when only the nodes (molecular components) and edges (interactions) are known (9). The most commonly applied logic is Boolean logic, where, for each node in the network, only two states are possible: “on” or “off.” A node in the “off” state does not indicate that the particular component is nonexistent in the system. Rather, it means that the component is not present at a high enough concentration to influence the state of the system it regulates (9). This concept is important when considering the relationship between the activating molecule and its target molecule. A target molecule will remain in the “off” state until the activating molecule reaches its functional activity concentration threshold (9). Similarly, a target molecule will remain in the “on” state until its inhibitor reaches its functional activity concentration threshold (9). Despite these nuances, some argue that the Boolean logic of “on” and “off” states in describing complex biological systems is too simplistic.
Özgür Sahin’s team at the German Cancer Research Center used a Boolean logic model of a cell line resistant to Trastuzumab, a therapy for metastatic breast cancer, to determine which target molecules to knockout to increase drug sensitivity in breast cancer patients (10). Hypothesizing that an escape from G1 cell cycle arrest conferred resistance, researchers constructed a Boolean logic network model of G1-S cell cycle regulation to overcome Trastuzumab resistance (10). They conducted simulations and validation experiments on proteins involved in receptor regulation and effectively identified a target protein in breast cancer cells (10).

Conclusion

Despite increasing investments in pharmaceutical research and development, scientists have struggled to produce effective treatments (1). Failure to identify effective novel therapies partly results from an inability to abandon reductionist approaches that emphasize single-target drug therapies. Systems biology recognizes the complexity of biological systems and uses this knowledge to better understand disease and drug mechanisms of action. It takes into account the networks in which target molecules play key regulatory roles and directs therapeutics to pursue network-based solutions.

The prospect of a systems biology approach in drug discovery seems promising in achieving a deeper understanding of biological systems. As systems biology advances, it will provide the tools needed to better predict outcomes of drug therapies at the molecular level.

CONTACT CHRISTINE PARK AT CHRISTINE.PARK.17@DARTMOUTH.EDU

References


Figure 4: Systems biology evaluates various network components to identify the best therapeutic targets.
Defining DVT and its Significance

Deep vein thrombosis (DVT), the third most common vascular disorder in the United States after heart disease and stroke, develops when blood clots form in a deep vein (1, 2). Deep veins received their name because they lie deep in reference to the skin compared with superficial veins, which lie just beneath the skin. Superficial veins drain into deep veins, which return blood to the right side of the heart. Deep veins drain into the right atrium and the pulmonary arteries, forming a channel that carries blood from the lower and upper extremities to the heart and lungs (1, 2).

Once DVT develops, blood clots may migrate to the right heart and pulmonary arteries. This migration may lead to pulmonary embolism (PE), which occurs when deep vein blood clots break off and migrate to the lung (3). DVT can also cause chronic leg pain and swelling. A more generic, inclusive term that describes both DVT and PE is venous thromboembolism (VTE), an often fatal disorder that kills 15 percent of patients within three months of its onset, making it deadlier than a heart attack (3).

DVT and PE have become increasingly prominent in the media, as high-profile athletes like Serena Williams, widely considered the best female tennis player in the world, and NBA stars Chris Bosh and Mirza Teletović have suffered from VTE. Studies also show that marathon runners are at a particularly high risk of VTE. Though athletes are generally perceived to be in prime health, recent studies indicate that athletes are at paradoxical risk for VTE due to their exposure to "classic" VTE risk factors (1, 4).

Virchow’s Triad and the Main Risk Factors for DVT

In the late 19th century, German physician Rudolf Virchow devised Virchow’s triad, which detailed the three classic risk factors of DVT (4). His triad consisted of 1) hypercoagulability, 2) circulatory stasis, and 3) vascular damage (4). Hypercoagulability, often due to genetic or acquired conditions, occurs when the blood has a greater tendency to clot. The next risk factor, circulatory stasis, is often due to patient immobility and occurs when the blood is stagnant. Vascular damage to the veins or arteries initiates the clotting process, also called the coagulation pathway (5). This pathway involves several coagulation proteins that are sequentially activated during vascular injury, which may result from an outside source like trauma or an inside source like atherosclerotic...
DARTMOUTH UNDERGRADUATE JOURNAL OF SCIENCE

plaque. These coagulation proteins eventually cause mature clots to deposit in veins (4, 5). Remarkably, more than one hundred years after Virchow developed his triad, these three risk factors are still critical in identifying VTE patients.

DVT Risk Factors in Athletes

Professional athletes frequently exhibit one or more of these risk factors for DVT, as their lifestyles render them particularly vulnerable to circulatory stasis and vascular damage. Athletes often experience circulatory stasis for long periods when taking flights to and from competitions. Two studies published by The Lancet and The Journal of the American Medical Association provided conflicting evidence regarding whether air travel poses a higher risk for DVT than ground travel (6, 7). Both studies have generated controversy regarding the hypothesis that the low pressure experienced during air travel increases DVT risk, but, nevertheless, the circulatory immobilization experienced during long flights increases DVT risk in athletes (6, 7). Elite athletes may also encounter vascular trauma more frequently than others. For instance, high-intensity sports such as running and high-speed cycling can induce repeated vascular trauma near the hip and the knee at the sites of flexion and extension (4).

Since circulatory immobility correlates with VTE, many people assume that exercise promotes circulatory flow, decreasing VTE risk. However, exercise, particularly strenuous exercise, is associated with increased clotting risk (4). One possible explanation for this phenomenon is rooted in evolution; when our ancestors exercised, it was often for survival reasons, and bleeding during a survival crisis was maladaptive. On the other hand, increased clotting during exercise and sympathetic activation of the autonomic nervous system was advantageous during crises, as it prevented excessive bleeding (8). Even today, the concentration of pro-thrombotic blood markers rises during exercise, while the fibrinolytic system, the body’s enzyme-based defense system that breaks down blood clots, is simultaneously activated during exercise (4).

Vascular injury remains a major risk factor for VTE and is the main reason for preventative measures such as pre-emptive blood thinners for hospital trauma patients (9). Trauma activates the coagulation pathway, causing not only vascular injury, but also circulatory immobilization and venous stasis and implicating all of Virchow’s triad (5, 9). Significant vascular trauma caused by overt injuries like broken bones increases the risk of DVT due to vascular damage and subsequent immobilization. Micro-traumas or injuries that are not evident could also initiate clotting mechanisms and cause VTE (4). Micro-traumas may result from repetitive low-intensity impacts such as cycling or long distance running. Similarly, sports that require a lot of arm movement, such as weightlifting, swimming, baseball, wrestling, and skeet shooting, often compress and inflict micro-trauma to upper extremity venous systems, potentially causing thrombosis of the axillary-subclavian vein, the main deep vein that drains blood from the arm (Fig. 2). No anatomic derangements are essential for this thrombosis to occur, but issues such as hypertrophied muscles at the thoracic outlet and extra-long transverse processes of the cervical spine and cervical ribs may facilitate axillary-subclavian vein compression and subsequent VTE.

Occasionally, athletes are also susceptible to hypercoagulability, possibly due to heritable genetic defects of clotting protein function, known as hereditary thrombophilia. Two of the most well-known markers for thrombophilia are protein C and S deficiencies, as these proteins inhibit the normal coagulation pathway (1). However, a mutation in factor V, a protein

Figure 2: The development of deep vein thrombosis in the leg.
essential for the coagulation pathway, causes protein C resistance and is the most common cause of VTE among caucasians. This mutation is known as the factor V Leiden mutation and accounts for approximately 20 percent of patients with a venous thrombosis (4). The second most common cause is a mutation in the prothrombin gene, which accounts for six percent of all VTE patients (4).

Hypercoagulability is not always due to hereditary conditions and can also be caused by an acquired condition such as post-exercise dehydration, which increases blood cell concentration. High hemoconcentration is a debatable risk factor for VTE, but hydration during exercise and long flights is still recommended (10). The environmental risk factors for VTE are old age, obesity, pregnancy, malignancy, trauma, surgery and smoking. In females, oral contraceptives, which contain hormones like estrogen, also increase hypercoagulability and VTE risk (3).

DVT Diagnosis and Treatment

The most common clinical presentation of a DVT patient is painful leg swelling (Fig. 5). Shortness of breath is also indicative of PE, although about a quarter of all DVT patients present with sudden collapse and death (11). Historically, PEs were discovered during autopsies because DVTs and PEs were both deadly and difficult to diagnose. Over the previous two decades, however, DVT has become easier to diagnose, leading to timely treatment and better survival rates. The standard treatments for DVT and PE are blood-thinning drugs that halt the growth of clots, allowing the body’s natural defense mechanisms to dissolve them (11). The faster the treatment is instituted, the better the chances of patient survival, meaning a quick diagnosis is paramount. Suspicion of the disease alone now causes physicians to treat with blood thinners while waiting for test results (3, 11).

Modern physicians screen for DVT using a blood test that indicates the presence of d-dimer, a protein fragment present in the blood after the patient’s intrinsic thrombolytic system degrades the clot (3, 11). If a d-dimer test is negative, the patient almost certainly does not have DVT, making it an excellent screening test. However, if a d-dimer test is positive, there is only a 50 percent chance that the patient actually has DVT, so other techniques are necessary to confirm DVT’s presence.

One such technique is Doppler ultrasound imaging of the legs, which more conclusively indicates the presence of deep leg vein thrombosis (3, 11). Doppler ultrasound sends sound waves into tissue and uses the reflected waves to produce images of the vasculature and any possible clots (11). Similarly, computerized tomography (CT) of the pulmonary arteries is a popular method for confirming the presence of a PE (11). In CT, X-ray images of the object are taken from multiple angles, and a computer then generates cross-sectional images of the scanned object.

In the past, PEs were diagnosed using an invasive pulmonary angiogram, a procedure in which a catheter is inserted into the pulmonary artery to inject contrast into the pulmonary circulation. This confirmatory technique was recently replaced with high resolution CT scans. After diagnosis, the patient is put on blood thinners for at least three months to avoid leg or lung clot growth, while the body’s inherent defense system slowly dissolves the clot. If an athlete has hereditary thrombophilia, he or she needs to take blood thinners permanently, while, if the athlete has a reversible cause, such as an injury, he or she only takes blood thinners for three months (3).
Decreasing the Incidence of DVT

To mitigate the risk of developing VTE, athletes should stay adequately hydrated to prevent their hemoconcentration from rising too high. For athletes, regular ambulation during long travel is recommended. They should also avoid maintaining the same position, such as crossed-legs, for prolonged periods, as this behavior could cause compression of veins that have already endured microtrauma. Additionally, wearing compression stockings during long flights could reduce DVT risk (Fig. 4). For athletes with family histories of VTE, screening for thrombophilia is recommended, and, for female athletes, contraceptives that do not contain estrogen might reduce VTE risk (10, 11).

Athletes with VTE may look fit, but, in actuality, they may not be completely healthy. If athletes present with leg swelling and shortness of breath, physicians should suspect VTE. The dissemination of education regarding DVT and PE to athletes and physicians may be the most promising measure, promoting earlier diagnosis and faster treatment.

CONTACT KEVIN KANG AT KEVIN.S.KANG.18@DARTMOUTH.EDU

References


Figure 4 (left): Serena Williams, one of the best female tennis players of all time, received emergency treatment in 2011 for a pulmonary embolism (PE).

Figure 5 (right): Chris Bosh, center for the Miami Heat, was forced to miss the end of the 2014-2015 NBA season due to deep vein thrombosis (DVT).
David and Goliath: The Future of Targeted Therapy in Cancer Treatment

EMILY XU, RALEIGH CHARTER HIGH SCHOOL

Cancer. More than one out of every three people will develop it in their lifetime, and it proves to be fatal in 7.6 million of those people per year (1). Ever since President Nixon officially declared the “war on cancer” 40 years ago, the National Cancer Institute has poured 90 billion dollars into research to curb this so-called “emperor of all maladies,” yet a cure still remains elusive. What may be the cause of this puzzling stalemate between nature and human intellect? The largest problem is simply one of language. When people refer to HIV, Lupus, or Polio, they are referring to a very specific spectrum of infective agents. However, cancer is not just one disease, rather it is hundreds, if not thousands, of diseases lumped together by common characteristics. Each individual type of cancer can acquire its oncogenic roots differently, and cancers that arise from the same cell type can be quite different. Cancer is also caused by the interaction of hundreds of mutations, so it is an extremely daunting task to try to find a way to identify and interfere with specific pathways. Additionally, tumors have the ability to metastasize, which makes them very difficult to treat. These metastases are completely different from the primary tumor due to independently accumulated mutations and, therefore, call for different treatments (2, 3). While some current cancer treatments have been effective in killing or removing cancer cells, such as surgery and radiation therapy, many times these treatments also end up killing healthy cells as well, forcing scientists to utilize the “therapeutic index,” which measures the chances of killing the tumor versus the chances of killing a normal cell (4).

In the face of all these formidable challenges, researchers have increasingly turned to a new interdisciplinary solution to treat cancer, targeted therapy. This new solution comes at the confluence of nanotechnology, biomolecular engineering, materials chemistry, and other fields to develop more effective drug delivery tools and techniques. The amalgamation and the interaction of these fields confer many advantages over traditional cancer treatment options. Because the size of the particles in these drug delivery systems is measured in nanometers, the toxicity of many of these targeted therapies decrease significantly (5). Introduction into the body also becomes a more manageable task, as the particles can be immediately absorbed into human tissues. It has even been reported that inhaled nanoparticles could reach the bloodstream from the lungs and make their way into other targeted sites, such as the liver or heart (5). Through the manipulation and selection of certain chemical materials, these particles can also be programmed to only release the drug when placed in a certain environment, making treatment more personalized and tumor specific (5).

Researchers from Brown University and the University of Rhode Island have utilized this strategy to develop a solution that takes advantage of the acidity of cancer cells. They have employed gold nanoparticles tethered to pH low-insertion peptides (pHiLPs), natural acid-seeking compounds, to hone in on the high acidity of the malignant cells (6). Then, they focus the energy of radiation into the area directly around the cancer cells. The radiation causes the nanoparticles to release a stream of electrons, inflicting damage on nearby cells (6). This nanoparticle harnesses the Auger effect, which allows gold atoms to interact with radio waves to release extra electrons due to the unique arrangement of electrons orbiting gold atoms (6). These effects allow for a very localized reaction (6). The team of researchers employed computer simulations and models to establish the quantitative details along with the convoluted calculations that this nanoparticle creation necessitated (6).

Physics and the properties of light and magnetism have also been enlisted in the fight against cancer, not only to create new nanoparticles that can carry drugs, but also to create ones that can enhance the crude, but widespread, treatments used today. Researchers have used magnetic nanoparticles to deliver heat directly to cancerous tumors. Applying heat directly to tumors allows for increased efficacy of radiotherapy and chemotherapy and reduces the needed dose (7). These magnetic nanoparticles made out of iron oxide are just a few tens of nanometers in diameter (7). They heat up when exposed to a powerful magnetic field and direct that heat to the tumors (7). Choosing the right kind of particle is crucial because different structures of nanoparticles deliver different doses of heat (7).

Another study at Oregon State University has developed a new system to selectively insert naphthalocyanine into cancer cells to allow for more accurate surgical removal of solid tumors and to eradicate any remaining cancer cells (8). Naphthalocyanine is special in that, when exposed to near-infrared light, it can make a cell glow as a guide for surgeons, allowing them to know where to cut (8). It can also produce reactive oxygen species that can kill the cancer cells directly (8). These two characteristics make it extremely attractive as a two-prong solution to cancer treatment. However, a couple of problems with this compound are that it is not water soluble and has a tendency to aggregate inside the body (8). As a result, researchers added a dendrimer, which is a special water-soluble nanoparticle, to encapsulate the naphthalocyanine (8). Dendrimers can slip easily into any tumor but will largely spare any healthy tissue (8). To append to the list of positive attributes of this solution, not only was the phototherapy successfully shown to destroy malignant tumors, but the laboratory mice also showed no apparent side effects or weight loss after the surgery (8).

Because of the thorough design that goes into these nanoparticles, different drugs can be combined to increase the efficacy of the treatments and to even target some of the harder to kill cancer cells.

A research team at the Cancer Science Institute of Singapore demonstrated the use of nanotechnology alongside existing chemotherapy drugs as agents against chemoresistant cancer stem cells (9). What makes cancer stem cells unique is their...
unusually high resistance to chemotherapy, which can lead to cancer recurrence (9). The nanoparticle developed was termed nanodiamond-epirubicin drug delivery complex (9). As the name suggests, the widely used chemotherapy drug epirubicin was attached to nanodiamonds, which are carbon structures with a diameter of around five nanometers (9). This combination of chemotherapy drugs with nanomaterials allowed for a broader range of protection in a capsule that is both safer and more effective. (9)

In a recently published article, a team of researchers used graphene strips to carry two anticancer drugs, namely TRAIL and doxorubicin, in a sequential manner to cancer cells, with each drug targeting the particular part of the cell where it has the highest efficacy (10). TRAIL, which is an anticancer protein, serves as an active targeting molecule that can bind directly to the cancer cell’s surface (10). Therefore, it is most effective when delivered to the external membrane of the cancer cell. On the other hand, doxorubicin works by intercalating DNA, and it most effective when delivered to the nucleus (10). The scientists utilized doxorubicin’s molecular structure to bind it to graphene and peptides to similarly bind TRAIL to the graphene (10). When this nanoparticle first comes into contact with a cancer cell, the receptors on the surface of the cell latch onto TRAIL (10). The cell then absorbs the remainder of the graphene with doxorubicin, beginning a process to trigger cell death (10).

This same group of researchers has also devised “nanodaisies” that have demonstrated promise to treat leukemia, breast, prostate, liver, ovarian, and brain cancers (11). These nanodaisies are made with a polymer called polyethylene glycol (PEG), which has long strands with shorter strands branching on either side (11). The researchers then utilized the hydrophilic properties of PEG and the hydrophobic properties of anticancer drugs camptothecin (CPT) and doxorubicin to create this daisy-shaped drug cocktail (11). The resulting particle is only 50 nanometers in diameter, which allows for easy delivery to the patient via injection (11). Once inside the patient, the nanodaisies are absorbed by cancer cells, and the two drugs attack the nucleus via different mechanisms, which allows for a proven increase in this approach’s efficacy (11).

New drug delivery systems are not only combining cancer drugs with other cancer drugs, but also with DNA. Jordan Green of The Johns Hopkins University School of Medicine’s Biomedical Engineering Department designed a nanoparticle delivery system that targets deadly brain gliomas in rat models, and it has been shown to significantly extend the lives of those treated rats (12). These nanoparticles are loaded with DNA encoding for the human simplex viruses thymidine kinase type 1 (HSVtk) protein (12). This gene produces an enzyme that converts ganciclovir, which has no effect on cancer cells in its natural state, into a potent destroyer of glioma cells (12). This team tested a variety of polymer structures for their ability to encapsulate and deliver DNA into the rat glioma cell lines. Their tests found that the polymer known as PBAE447 was the most efficient in delivering the gene and was shown to be 100 percent effective in killing both of the glioma cell lines when combined with ganciclovir (12). Similar results were found using live rat models as well (12). This system avoids the problems associated with viral delivery methods such as toxicity, a triggered immune response against the virus, and the possibility for the virus itself to induce tumorigenesis. It is far more effective than traditional drug and radiation therapies (12).

However, these improvements do not come without a cost. It is predicted that these new treatments will have a significant cost in research and production. Others worry that like many antibiotic resistant bacteria, cancer cells will also become resistant to these targeted therapies (13). In the future, this would necessitate the use of even harsher drugs to treat patients for the same diseases (13). Another huge concern that critics of targeted therapy have is the permanence of the nanoparticles. While some have the ability degrade and dissolve easily, others that aren’t degradable or soluble can accrue in the human body for a sustained period of time.

Despite still being in a state of infancy, targeted therapy is making significant strides in prevention, detection, and eradication of what has been largely known throughout the course of human history as an incurable disease. With new discoveries and advances in nanotechnology, biomedical engineering, and chemistry occurring daily, this interdisciplinary approach is quickly gaining more and more traction from oncologists and researchers worldwide. To neglect the sophisticated advantages of targeted therapy is to forgo one of the most promising interdisciplinary innovations in cancer treatment that may finally make a cure for cancer within reach.

CONTACT EMILY XU AT EXU@RALEIGHCHARTERHS.ORG

References

Bioinformatics: Piercing and Pinpointing Moonlighting Proteins and Beyond

WILLIAM (WOOSUNG) JUNG, KOREA INTERNATIONAL SCHOOL

Biologists build upon prior knowledge in order to achieve great scientific discoveries and developments. The laboratory was the crucible of scientific inquiry, an arena where hypotheses and theories were tested and answered. However, since the arrival of the computer, another field has evolved that bolsters the abilities of scientists in both scope and speed: bioinformatics. According to Luscombe, Greenbaum, and Gerstein (2001), "bioinformatics [conceptualizes] biology in terms of macromolecules...and then applies 'informatics' techniques (derived from disciplines such as applied math, computer science, and statistics) to understand and organize the information associated with these molecules on a large scale" (1). What once took scientists grueling hours and massive funding may now be achieved through Perl coding and C++ programming. However, computer programming itself is only a tool that can be tapped by scientists. Tools may till the earth or chart the stars, but they also require human minds to guide them – to maneuver around rocks and to look beyond the sky. Evaluating the fields of biology and computer science, this paper seeks to probe the individual strengths and limitations of both fields as they aid in the research of protein multitasking, also known as moonlighting. Following their individual analysis, this paper discusses why bioinformatics offers one of the greatest opportunities for piercing the moonlighting mystery through an interdisciplinary approach. First, however, we shall discuss the meaning and significance of moonlighting proteins.

Moonlighting proteins are a class of multifunctional proteins: the single protein is "capable of performing multiple physiologically relevant biochemical or biophysical functions that are not due to gene fusions, multiple RNA splice variants, or pleiotropic effects" (2). One way in which moonlighting proteins affect how our cells function can be observed when an enzyme moonlights in response to changes in cellular conditions, prompting changes in the regulation of a biochemical pathway such as glycolysis or, in plants, photosynthesis. By changing the biochemical pathway, moonlighting alters the process of catalyzing substrates and the products formed through catalysis. But the roles of these proteins are not limited to regulating the translation and transcription of DNA and RNA; in fact, moonlighting proteins have expansive impacts on the broad spectrum of our biological world. An enzyme in glycolysis, phosphoglucone isomerase, also serves as a cytokine involved in metastasis in breast cancer tumors (3). Examining these moonlighting proteins may thus enable our foray into the territory of uncharted diseases, as well as corresponding vaccines and remedies (4). By understanding the various phases of moonlighting an enzyme undergoes, doctors could intervene in the metastasis of cancerous growths, dropping mortality rates to an unprecedented degree. Because of the field's current obscurity, there is still a vast reservoir of potential medical and biological knowledge yet to be tapped. To sift through the nebulous clouds surrounding moonlighting proteins, however, biologists face the daunting task of parsing through massive quantities of data, "usually, [moonlighting proteins] are revealed experimentally by serendipity, and the proteins described probably represent just the tip of the iceberg" (5). But luck cannot be deemed a solid foundation for scientific inquiry. Instead, as we learn more about moonlighting proteins, scientists are able to identify characteristics and their patterns with growing certainty, aiding in further discoveries. While this task of identifying unknown attributes of moonlighting is a major challenge for the scientific and technological community, as discussed later, the whole process begins with biology.

Biologists around the world are akin to librarians in the study of our living world. Their chief onus, in the scope of this essay, falls under ontology, the specification and description of a natural object's existence. In this way, Felton (2002) points out that biologists are responsible for cataloging and partitioning their observations for further research and others' reference: "In the past, biologists and chemists made improvements in their laboratory methods, making dramatic leaps in understanding, but also enabling other scientists to discover new information based on the improved techniques. Take, for example, the polymerase chain reaction. This discovery cobbled together existing information, but it became a tool that thousands of scientists have used to further their own study" (6).

Unfortunately, this responsibility carries a heavy burden. One of the most significant hindrances faced by biologists and chemists who seek to build upon the findings of others is semantic variance – the unavoidable fact that language and descriptions are subjective and, thereby, are not always perfectly accurate (7). This inaccuracy compounds the difficulty faced when attempting to research a subject as massive and nebulous as moonlighting proteins because it adds uncertainty to every ontological definition of a protein. Even so, it is that uncertainty that compels biologists to review and discover the unheard-of attributes of cells and their respective functions (5). Especially since the completion of the Human Genome Project in 2001, their next leap would be either to develop "a novel program" or to find "a unique way to couple existing program[s]" instead of holding on to the growingly inefficient wet laboratory style (6). In this vein, it is certain that "a significant part of the future of biology lies with computer methods" (6).

Through observations made by biologists who seek moonlighting proteins, hints and clues can be discovered that may never have been identified precisely because no other scientist was specifically looking for them. By broadening their focus, biologists sacrifice time in favor of scope; but now there is a tool that can help them save that most precious resource. Computer programming and coding are not restricted
to contributing to any particular field; they are a powerful multipurpose tools for the processing and analysis of vast quantities of data. A specific tool utilized by programmers with particular relevance to the perception of moonlighting proteins is called “data mining.” Data mining is the identification of statistical patterns, predictive models, and hidden relationships, usually among massive amounts of data. This tool was primarily used for economical analysis, but many have posited that it could be applied to the field of biology to great effect. Indeed, the successes of this union between computational data analysis and biological study created the field of bioinformatics (1). The advantage of applying computer programming and coding to biomedical research through data mining is not only that it saves immense amounts of time, but also that it may help standardize the ontological descriptions of the human genome, further helping to advance the identification and understanding of moonlighting proteins. Yet the sole application of bioinformatics itself is inherently dependent upon the expertise of those who analyze the data it generates, which necessitates the presence of an individual or individuals who are well-versed in programming as well as biological research (8). Indeed, one of the core reasons it is so important to unify biological and programming knowledge is that there is a finite amount of data that can be appraised by biologists themselves. As a result, some data sets that are processed may need to be discarded in order to allow memory and attention to be allocated to new, more promising inquiries. Unfortunately, computer programmers alone do not have the required expertise and scientific knowledge to make such decisions independently, thus necessitating the active engagement of biologists in the interdisciplinary field of bioinformatics.

By combining the field of biology with information technology, research groups have already begun compiling databases of moonlighting proteins with standardized descriptions and identifying attributes for the biomedical community to appraise (9). The biologist’s “encountered the difficulty of collecting examples of [multitasking] proteins because of the lack of a broad database, so the effort to gather the examples was often one of the main challenges” (9). As a result of their compilation, the research group realized that the existence of moonlighting proteins may be even more significant than previously speculated (9). What if, for instance, the multitasking was not restricted to a pair of functions? Hernández and his team posited that there may be three or more moonlighting functions which, without their database or one like it, would require excessive time and no small luck to discover (9). Called “MultiTaskProtDB,” their database lists over 288 moonlighting proteins, their NCBI and UniProt accession numbers, canonical and additional biological functions, monomeric/oligomeric states, and bibliographic references so that others can follow research that supports the database (9).

The system itself was created using MySQL software to form the database, while coding was written in PHP to help researchers refine their searches for particular moonlighting proteins within the database (9). At the same time, “MoonProt,” another database that also lists the known moonlighting proteins in a system was created (10). At first the creation of two databases may seem redundant, but there is certainly a benefit to this overlap. By creating these interrelated databases, Mani and Hernández’s projects offer a chance to compare and contrast the ontological accuracy of their proteins to form standardized profiles for more than 200 proteins (9, 10). It also provides constant comparison between the two databases, allowing bioinformaticians to perceive what technology, functions, and biological queries are most promising or, in some cases, need to be improved for the field to advance.

The value of accuracy in both medical and scientific communities cannot be overstated, yet there are still challenges that must be faced by each field, whether together or alone. For biologists, the time-consuming and challenging process of identifying moonlighting indicators will be augmented by the raw processing power of modern coding. However, even with state-of-the-art technology, databases are incapable of brute forcing through million lines of inconsequential data due to physical memory and storage constraints. To improve the quality of data and increase efficiency, cooperation of biologists is crucial in assisting in the perception of new moonlighting identifiers and proteins. Bioinformatics is not only one of the most groundbreaking interdisciplinary approaches to decrypting the human genome; it also empowers scientists, researchers, and doctors with one of the most important abilities in their professions: being able to inquire and seek knowledge. Through these databases, hypotheses can be tested and evaluated with nearly infinite scale and in unprecedented brevity. Representing the union of scientific minds and the flawless retention and lightning-fast accessibility of virtual storage, bioinformatics is a wonder to behold and a key to the future of modern scientific and medical progress.

CONTACT WILLIAM JUNG AT WOOSUNG.WILLIAM.JUNG@GMAIL.COM

References

BOOK ANNOUNCEMENT

What Every Science Student Should Know

YOO JUNG KIM ’14 AND ANDREW H. ZUREICK ’13

Four recent Dartmouth alumni, including three past presidents and editors-in-chief of the Dartmouth Undergraduate Journal of Science (DUJS), recently came together to publish a book entitled What Every Science Student Should Know. We are pleased to be able to announce our book’s release in this Winter 2016 issue of the DUJS.

What Every Science Student Should Know aims to serve as a mentor for the aspiring STEM (science, technology, engineering, and mathematics) student. This book is a distillation of our own experiences as recent science graduates, bolstered by years of research and 100+ interviews with successful scientists and other science students. Like an experienced lab partner or academic advisor, the book points out the challenges students commonly face when pursuing a science major while providing encouragement at every step of the process. Chapters cover the entire college experience, including choosing a major, mastering study skills and time management, conducting scientific research, finding a job, applying to graduate and professional schools, and, most importantly, how to foster and maintain a lifelong love of science. What Every Science Student Should Know is our advice on how to excel as both a student and a scientist.

Co-authors Justin L. Bauer ’12, Yoo Jung Kim ’14, Andrew H. Zureick ’13, and Daniel K. Lee ’13; University of Chicago Press editor Christie Henry ’91; and many other Dartmouth College students, alumni, professors, programs, and organizations have been an essential part of the publication of What Every Science Student Should Know. But no other organization has been as integral to our book as the Dartmouth Undergraduate Journal of Science (DUJS). Our collective experiences working for the DUJS gave us the technical skills and the confidence to tackle the arduous task of a composing a 256-page book manuscript.

What Every Science Student Should Know is now available online for pre-order, and it will be officially published and shipped in May 2016. If you place an order on the University of Chicago Press’s website using the promotional code “PRSCISTU20” by 4/15/2016, you can receive 20% off the list price (http://www.press.uchicago.edu/ucp/books/book/chicago/W/bo19167321.html).

Additionally, the book is available for pre-order on Amazon (http://amzn.com/022619888X) and other major book retailer online stores.
Evaluating Pathological Angiogenesis in Wild Type vs. Slug Knockout Mice

ARVIND SURESH1

Abstract
Angiogenesis, or the growth of new blood vessels from preexisting blood vessels, is required for both developmental and pathological processes, such as tumorigenesis. Slug, a member of the Snail family of transcription factors, is an important promoter of epithelial-to-mesenchymal transitions, a process that is critical for the initiation of cancer metastasis. Through an in vitro study, Slug has recently been shown to induce endothelial cell (EC) sprouting angiogenesis. Preliminary data using the Matrigel angiogenesis assay suggest that Slug may also influence pathological angiogenesis in vivo. However, the exact role Slug plays during this process is still unclear. In this study, we compared the level of angiogenesis in Matrigel plugs from both wild type and Slug knockout mice using image analysis. We observed a loss of EC recruitment and vessel formation in Slug knockout mice. The results of this study shed light on potential mechanisms by which Slug affects pathological angiogenesis. This knowledge could be used to develop future cancer treatments.

Introduction
Angiogenesis is a multi-step process essential for normal growth and development, as it enables the body to supply cells and tissues with oxygen and nutrients (1). It is defined as the physiological process by which new blood vessels form from preexisting blood vessels, which are composed of endothelial cells (ECs). Angiogenesis begins when the ECs composing the vessel wall lose their basal-apical polarity and start degrading the basement membrane. Following this step, the nascent sprout, guided by a highly migratory “tip cell” and trailed by proliferating “stalk cells,” extends into the extracellular matrix and forms lumens that can eventually undergo anastomosis with other existing vessels to form new mature networks (2). In normal developmental growth, angiogenesis is critical for wound repair and uterine lining development during pregnancy (3). At the same time, angiogenesis has been found to treat ischemic heart disease, providing a steady blood supply to tissues affected by plaque deposits in coronary arteries (4). However, angiogenesis has also been implicated in pathological processes, such as tumor growth and metastasis (2). Stimulated by pro-angiogenic factors such as vascular endothelial growth factor (VEGF), the excessive vessel growth around tumor cells not only supplies them with the nutrients necessary for rapid growth, but also with channels for cancer cells to invade surrounding tissues and metastasize to new sites (2).

Slug is a member of the Snail family of zinc-finger transcription factors involved in embryonic development processes such as mesoderm, neural crest, and heart cushion formation (5, 6). It also combats apoptosis and promotes epithelial-to-mesenchymal transitions, a process critical for the cancer metastasis initiation (7). Although the developmental role of epithelial Slug expression has been well established, studies concerning the effect of autonomous expression of Slug in ECs on cell behavior are sparse. With respect to Slug’s relationship to angiogenesis, our lab has recently shown that Slug is also crucial to sprouting angiogenesis in vitro (7). By utilizing a three-dimensional fibrin bead angiogenesis assay, we demonstrated that Slug overexpression in ECs led to greater angiogenic sprouting when compared to a control group with GFP expression alone, providing proof that Slug promotes angiogenesis. However, due to the differences between in vitro and in vivo conditions, it was necessary to test if and how Slug influences pathological angiogenesis in vivo. To do this, we performed the Matrigel angiogenesis assay on wild type and Slug knockout transgenic mice (8, 9). By injecting mice with Matrigel, a basement membrane protein extract from mouse sarcoma that helps mimic the tumor environment, we observed that Matrigel plugs from the Slug knockout (Slug-) mice have a much lower amount of blood in comparison to the wild type (Slug+) control group (10). This led us to wonder why there is a difference between the two genotypes and the exact role Slug plays in this process. Here, we provide evidence that Slug is an important mediator of EC recruitment and vessel formation in pathological settings. Curiously, we find that Slug also influences stromal cell density and vessel density, and may even play varying roles between gender types. As a whole, these studies suggest the impact of Slug deficiency in tumor angiogenesis.

Materials and Methods

Microscope Slide Preparation and Imaging
Images of Matrigel plugs were taken from three wild type (labeled D, E, F) and three Slug knockout brown mice (labeled A, B, C). In each set, two of the mice were male, while one was female. Slides were sectioned and stained with CD31-1 for ECs from the top, middle, and bottom of the plug (referenced by 1, 2, and 3). Using an Olympus microscope with a Color Mosaic camera and SPOT imaging software, five to 10 specific areas of each plug were captured at 100x magnification (Fig. 1a). A second set of Matrigel CD31-stained sections were obtained from the same plugs, and five randomized regions of the Matrigel in contact with native tissue were captured again at 100x magnification for each slide.

1 Molecular Biology & Biochemistry, University of California, Irvine, CA
Research Advisor: Nan Wu, Molecular Biology & Biochemistry, University of California, Irvine, CA
Faculty Editor: Christopher C.W. Hughes, Department Chair, Molecular Biology & Biochemistry, University of California, Irvine, CA
Using National Institutes of Health (NIH) Image J software, images were initially processed by drawing boundaries and identifying regions of 50, 150, and 250 micrometers from the edge of the native murine tissues (Fig. 1b). Only the vessels found in the regions beyond 50 microns from the native tissue boundary were considered newly formed (11). Initially, percentage EC area was calculated by identifying the area in microns$^2$ that were stained positive for CD31 in regions between 50-150 microns, 150-250 microns, and 250+ microns from the native tissue. The same was then done for vessel areas, which were identified as collections of ECs and/or round cross-sections of hollow tubes outlined by ECs. Using a Cell Counter plugin in Image J, both stromal cells whose nuclei are stained blue by a Haematoxylin counterstain and individual vessels were counted in each of those same regions (Fig. 1c). Lastly, based on their diameter, circular vessels were classified as large ($\geq$ 35 microns), medium (19-34 microns) and small ($\leq$ 18 microns). For the second set of sections, the same image analysis strategies were applied to randomized regions, but only the percentage EC area for the entire Matrigel area beyond the first 50 microns was calculated.

Data Analysis and Statistical Analysis

Percentage EC area and percentage vessel area were calculated by dividing EC area and vessel area by total Matrigel area. Vessel density and stromal cell density were calculated in units of count per 100 micron$^2$. Averages of each region were initially calculated for each section. Averages of all sections from each mouse were then taken and grouped by genotype. To compare the Slug knockout to the wild type control group, averages of the three mice in each group were taken and reported through a clustered column bar graph with standard error of the mean (SEM). During the second trial, averages of the values were used to compare several different categorical variables, including mouse genotype, mouse gender, and the combination of the two. Statistical significance was evaluated using a two-tailed T-test with equal variance; p-values of < 0.05 were considered statistically significant.

Results

 Slug Deficiency Affects Endothelial Cell and Stromal Cell Invasion

The presence of both ECs and stromal cells in the Matrigel was greatly reduced in the Slug knockout mice (Fig. 2). In general, percentage EC area experienced a greater reduction in comparison to stromal cell density, while the respective magnitudes of reduction between the individual regions for both graphs carried similar trends (Fig. 2a & b). Such reductions in percentage area and density could be attributed to EC loss and stromal cell invasion at these regions. With respect to percentage EC area, a general trend of reduced invasion was present at all distances from the native tissue boundary in the Slug-deficient mice. However, the greatest reduction between the wild type and knockout mice was observed in the region farthest from the native boundary beyond 250 microns, while the reduction in total area appeared to be an average of those at the three individual regions. At the same time, the large reduction between the genotypes after 250 microns implied that Slug promoted long-distance invasion and migration of ECs.

When examining stromal cell density at 50-150 micron, 150-250 micron, and 250+ micron regions, the Slug knockout mice experienced reductions of 33.2 percent, 47.4 percent, and 72.8 percent, respectively, in comparison to the overall total, which also retained an average of the reductions at the individual regions. In both cases, the regions 250+ microns from the native tissue had the highest standard error of the mean (SEM) of 1.0226 for EC area and 0.0049 for stromal cell density, as well as the greatest difference in magnitude between the two genotypes, with reductions of 94.5 percent and

Figure 1: Image Analysis of Matrigel Plugs. (A) Montage of one full Matrigel plug. Matrigel is in the center, which is surrounded by native tissue. (B) Image demonstrates the method used to draw boundary lines at 50, 150, and 250 micrometers. (C) Sample image demonstrates the identification of endothelial cells, vessels, and stromal cells in each plug.
72.8 percent. A visual representation of this reduction between the phenotypes can be observed through representative images of both the wild type and Slug knockout mice (Fig. 4a & b). After conducting a second trial with the same set of specific images for EC area (Fig. 5), a similar trend between the genotypes was observed. However, there appeared to be a reduction in variability, represented by the SEM, of 57.3 percent and 62.5 percent, further emphasizing the trend towards a reduction in EC recruitment in knockout mice. This also resulted in a decrease of the average p-values for the data set by approximately four percent, proving that the lack of statistical significance between the genotypes could be corrected with repetitive trials and a larger sample size. These data sets of data indicate that Slug deficiency reduces percentage EC area and stromal cell density at all distances from the native tissue boundary, affecting the invasion and migration of cells important to angiogenesis.

**Slug Deficiency Affects Vessel and Tubule Density and Size**

After analyzing data concerning endothelial and stromal cells, it was necessary to confirm whether Slug affected EC recruitment, the arrangement of cells into vessels and/or tubules, or both. Both

![Graph A](image1.png)  
**Figure 2:** EC and Stromal Cell Invasion in Matrigel Plugs. (A) Bar graph comparison between the WT and Slug KO mice examines percentage of EC area at three different regions and measures a total for the entire Matrigel plug. (B) Bar graph measures percentage stromal cell density/100 micron².

![Graph B](image2.png)  
![Graph C](image3.png)  
**Figure 3:** Analysis of Vessel Size and Density in Matrigel Plugs. (A) Bar graph comparison between the WT and Slug KO mice examines percentage vessel area at three different regions and measures a total for the entire Matrigel plug. (B) Bar graph measures vessel density/100 micron². (C) Bar graph measures the number of vessels for three vessel size categories.

the graphs of vessel area and vessel density indicated very similar trends between the wild type and Slug knockout mice. Again, the bars at 250+ microns had the largest change in magnitude, with a 98 percent reduction in the Slug knockout mice, while the overall reductions for each graph were 83 percent and 79 percent for vessel area and vessel density, respectively. However, when these vessels were categorized, there was a greater reduction when comparing the two genotypes in small and large sized vessels of 42.9 percent and 63.3 percent, respectively, in comparison to medium sized vessels. The number of medium-sized vessels was not significantly affected, however, with a reduction of only 28.6 percent. In addition, both genotypes carried an average of 41 percent more small vessels than medium and large vessels. Therefore, although the trend in vessel size ratio remained similar between the genotypes, the trend in vessel count did differ (Fig. 3a-c). This may have resulted from Slug-induced proliferation of ECs that contribute to small vessel and tubule formation. Alternatively, this discrepancy could be attributed to a Slug deficiency that prevented an increase in vessel diameter, thereby revealing Slug’s role in vessel expansion. Collectively, in the Slug knockout mice, there were fewer vessels and a smaller percentage of vessel area inside the Matrigel.

**Categorical Study Reveals Possible Genotype and Gender Relationships**

The secondary trial (Fig. 5) also demonstrated that the effects of gender, genotype and combinations of the two could be compared to establish general correlations and associations. From the two bars at 50-150 microns on the mouse gender graph (Fig. 5b), it can be seen
Figure 4: Comparison of Representative Matrigel Images. (A) Wild type representative image contains large and medium vessels, groups of individual endothelial cells, and many stromal cells that surround the ECs. (B) Slug knockout representative image lacks both vessels and endothelial cells, but contains large numbers of stromal cells at all distances from the native tissue boundary.

Figure 5: Secondary Trial and Categorical Comparison between Gender and Genotype. (A) Bar graph comparison between the WT and Slug KO mice examines percentage EC area at three different regions and measures a total for the entire Matrigel plug. (B) Bar graph measures same criteria between males and females. (C) Bar graph compares female wild type mice with female knockout mice. (D) Bar graph compares male wild type mice with male knockout mice.

that there was a statistically significant difference between the two bars with a p-value of only 0.048. As a result, it is possible that gender affected EC recruitment, especially when coupled with genotype. This is demonstrated in Fig. 5c, in which there is a significant difference between the total bars for the graph of percentage EC area when comparing female wild type mice to female knockout mice. These conclusions may be influential when studying the combined effects of gender and genotype with a larger sample size of mice.

Randomization Proves Consistency of Matrigel Plug

After conducting a randomized study using Matrigel slides to calculate percentage EC area (Fig. 6), a striking similarity could be observed between the original data set using specific area regions and the randomized study results. In both instances, randomizing the areas of sampling did not alter the trends established in the first and second trials regarding EC area. However, since a Matrigel plug affected by extraneous factors such as location could not be expected to remain consistent throughout, there was larger variability, represented through SEM bars and higher p-values. The results of randomization were successful in corroborating the results of previous trials throughout the experiment.

Discussion & Future Directions

Recent studies have revealed the new role of the Slug
transcription factor in angiogenesis, but the exact mechanism by which Slug regulates EC sprouting angiogenesis has yet to be elucidated (7). In this study, we analyzed the Matrigel plugs of wild type and Slug knockout mice for the percentage EC area, stromal cell density, vessel density, and vessel size. From the results, we concluded that there was a trend towards a loss of EC recruitment and vessel formation in the Slug knockout plugs. This conclusion might also be expanded to include stromal cells. Despite initial hypotheses that Slug would only affect ECs, it seemed that Slug also affected stromal cell invasion from surrounding native tissue. This finding could either be a direct result of Slug expression in stromal cells or a secondary effect of Slug expression in ECs that in turn affected EC-stromal cell interaction. Regardless, it is possible that Slug plays an important role in the communication among cells involved in angiogenesis. Moreover, the greater reduction in vessel formation as compared to EC recruitment in Slug knockout mice suggested that Slug may play a greater role in vessel formation. However, this finding could also be a compounded effect from EC loss, which could hinder vessel formation.

A study of vessel categorization showed that the greatest differences between the two genotypes were largest at the ends of the size spectrum. Without further experimentation, a conclusive relationship between vessel size and percentage vessel density cannot be drawn. Additionally, a trend towards greater EC recruitment was present in the female gender upon comparison of percentage EC area between mice of both genotypes (Slug+/ Slug-) and genders (M/F). At the same time, a lack of statistical significance most likely resulted from a small sample size and large variations within each sample size. For future experimentation, it might be best to repeat the experiment with larger sample sizes and introduce statistical blocking to further reduce error. For greater accuracy, the use of color contrast technologies and computerized randomization may be used to quantify data and account for human error, validating the conclusions made in this investigation.
The Effects of the Mediterranean Diet on Symptomatic Patients with Gastroparesis

SARAH BESSEN

Abstract

Gastroparesis (GP) is a chronic neuromuscular disorder of the upper gastrointestinal tract defined by dyspeptic symptoms in the absence of physical blockage. Current treatment, although helpful, results in continued diet alterations and inflammation. Recent studies have acknowledged the Mediterranean Diet (MD) for reducing cellular and circulating inflammatory biomarkers, making its potential in GP treatment of particular interest. We examined the effects of the MD on symptomatic adults diagnosed with GP, evaluating food diaries and C-reactive protein and pro- and anti-inflammatory cytokine levels in the blood before and after six weeks on the MD. Of the subjects who completed the study, there was a significant difference seen in the ratio of saturated fat to total fat intake in the first week, and tumor necrosis factor-α and interleukin-8 biomarker levels significantly decreased in evaluable subjects six weeks following MD initiation. Two of the five patients lost weight (10-25 pounds), and two others were able to intentionally gain weight with increased food tolerance on the MD while decreasing saturated fat to total fat ratio. Three of the five subjects had diabetes as the cause of GP, and each of these subjects reported a decreased insulin requirement. The most dramatic difference was observed in quality of life (QOL) descriptions. As with all studies involving change of diet, compliance levels were low, with the best results observed within the first week. With increased glycemic control and quality of life reported, efforts to change poorly established eating habits in this population might be warranted.

Introduction

Gastroparesis (GP) is a chronic neuromuscular disorder of the upper gastrointestinal tract that affects at least four million people in the United States and is defined by dyspeptic symptoms and delayed gastric emptying in the absence of physical blockage (1, 2). In general, symptom duration is often a third factor added to the definition, as many acute illnesses or abdominal operations can temporarily impair stomach function. Lasting subjective symptoms include nausea, vomiting, bloating, early satiety and fullness, and pain. Ninety percent of GP patient experience nausea, 84 percent experience vomiting, 75 percent experience bloating, and 60 percent experience early satiety (3). As many of these symptoms result from eating, GP is also linked to nutritional and metabolic disorders (2). Patients with GP often suppress eating, resulting in the insufficient intake of dietary calories, vitamins, and minerals (4).

Symptoms of GP can range from mild to completely disabling, meaning treatment can range from minor dietary modification to frequent hospitalization and enteral and intravenous nutrition, meaning quality of life can be affected in several different ways (5). Several studies have found that GP significantly impairs quality of life, as patients have significantly lower scores on tests that assess physical and mental or emotional function (3). Because it is difficult and often impossible to alter the underlying processes that cause GP, current treatment largely depends on presented symptoms, with the goal of improving overall quality of life in spite of a sometimes disabling disorder (6). GP can vary tremendously from one individual to another in terms severity, symptoms, and etiologies, and, as such, there are a variety of treatment options that are considered for GP patients. For example, because gastric emptying even in healthy individuals is affected by nutrient density and food consistency, dietary modifications are important to limit symptoms and prevent nutritional deficiencies.

Dietary management is especially important for diabetic patients, who must balance the relationship among blood sugars, insulin therapy, and gastric emptying. Prokinetic drugs, namely erythromycin, metoclopramide, and ghrelin, are often the first choice of therapy and aim to treat GP by accelerating gastric emptying (2). In various cases, antiemetic therapy is used to target symptoms such as nausea and vomiting that affect a patient’s ability to tolerate oral food intake. Antidepressants are commonly used to lessen nausea and pain, as well as to address the high prevalence of anxiety and depression, as well as the role of emotion on symptom severity (2).

In other cases, patients turn to surgical options. Some operations target a complete removal of the stomach, while other surgeons choose to perform pyloroplasties in order to improve gastric emptying. Because of the known relationship between electrical activity of the gastric muscle and gastric emptying and symptoms, gastric electrical stimulation has been used to treat GP in the event that other therapies do not provide adequate relief. The only current FDA approved option Enterra™ Therapy by Medtronic, which electrically stimulates the lower stomach and has been successful for many patients. A Medtronic sponsored clinical trial compared 12-month post-implant to baseline data and found that the therapy may be used to reduce symptoms of nausea and vomiting in patients that could not be
sufficiently controlled with standard therapies. In this study, 93 percent of patients experienced a reduction of vomiting episodes greater than 50 percent, and 53 percent of patients experienced a reduction of vomiting episodes greater than 80 percent at 12 months versus baseline (7).

While any of the numerous therapies alone or in conjunction with one another may lessen symptoms, the limited efficacy of the therapies often leaves patients frustrated and affects quality of life. Widespread recognition of the role of inflammatory processes and the overall presence of inflammation in patients with GP provides a potential site of discussion, especially in conjunction with emerging literature on the popular Mediterranean Diet (MD) as an anti-inflammatory diet. The MD has long been lauded for reducing cardiovascular risk, and recent studies have recognized the anti-inflammatory impacts of MD for reducing cellular and circulating inflammatory biomarkers such as serum C-reactive protein, interleukin-6, and endothelial and monocyctary adhesion molecules (8, 9, 10). Accordingly, some clinicians have proposed that the MD could alleviate the symptoms of GP; in this investigation, I will evaluate this prospect.

Methods

Inclusion criteria included patients diagnosed with GP by a gastric emptying exam who were also above the age of 18. Participants were asked to resume normal eating habits and record daily for a week everything they consumed in a food diary provided by investigators, along with their weight, weekly expenses, nutritional supplements, and any physical or emotional observations. Patients were also asked to complete a Medtronic Patient Diary for GP symptoms provided by investigators in order to record symptoms specific to GP, including nausea, pain, vomiting, and satiety. Following one week of normal eating habits, patients were interviewed regarding their current eating habits and quality of life, had their blood drawn, and had their heights and weight measured. The data collected from this week represented control data.

Participants were instructed on the principles of the MD and asked to adjust their diet accordingly. Participants were instructed complete food diaries and Medtronic Patient Diary for GP symptoms for six weeks. At the end of the six weeks, participants had their blood drawn and were interviewed using the same questions asked in the preliminary interview.

Food diaries were analyzed for total calories, proteins, carbohydrates, total fat, saturated fat, polyunsaturated fat, and monounsaturated fat. Proteins, carbohydrates, and fats were measured in grams, and values were taken from the popular food database Supertracker (11). In the event that the food item was not listed on Supertracker, investigators used another database called FatSecret (12). Blood drawn at the beginning and end of the investigation was analyzed for C-reactive protein (CRP) levels, a measure of inflammatory response. CRP is secreted by the liver in response to various inflammatory cytokines and is widely used to monitor inflammatory states (13). Blood was also analyzed for interleukin-6 (IL-6), interleukin-1β (IL-1β), interleukin-10 (IL-10), interleukin-6 (IL-6), interleukin-18 (IL-18), and tumor necrosis factor-α (TNF). Descriptive statistics were used to determine any significant differences in nutrition patterns and cytokine levels. Prior to analysis, patient information was de-identified, and patients were numbered from one to 11 based on their start date for the investigation.

Results

Eleven subjects began the study, and three dropped out. Of the eight remaining subjects, five were compliant with the diet and subsequently used for analysis (Patients 1, 3, 4, 8, and 9). There was a significant difference observed in the ratio of saturated fat to total fat in the first week. There were no differences in CRP, IL-1β, IL-10, and IL-6 levels, however, TNF-α levels significantly decreased in evaluable subjects six weeks following initiation of the MD. Two of the five patients lost weight (10-25 pounds). Two other subjects were able to intentionally gain weight with increased food tolerance in the MD, while decreasing their saturated fat to total fat ratio. Total saturated fat intake decreased as reported by the subjects, but this decrease was not significant. Three of these five
patients had diabetes as the cause of GP, and each diabetic patient had a gastric stimulator placed in the past. All three diabetic patients reported decreased insulin requirements.

Three of five participants reported that they disliked completing a food diary. Patient 8 reported disliking the process of recording food consumption and felt very self-conscious. Patient 8 explained that recording food sometimes influenced eating choices because it felt as though there were “other eyes on [the patient]”. Patient 3 similarly disliked recording a food diary, and Patient 4 agreed but felt that the task had become routine. While patients 3, 4, and 8 each expressed that they disliked maintaining a food diary, each also described the process as helpful. Patient 4 explained that the process created an awareness of foods he/she needed to consume for a more complete and healthy diet, as well as foods that he/she should avoid that might trigger symptoms. Patient 9, who succeeded with his/her goals of gaining weight, also expressed that the process of recording his/her diet had been encouraging, as it showed what he/she was actually eating, made him/her aware of progress, and showed him/her impressed at “how healthy [he/she]” was while on the MD.

Discussion

That the only significant difference in terms of food consumption was in the ratio of saturated fat to total fat in the first week mirrors the majority of literature regarding diet change and subjects’ likelihood to make long-lasting changes. Behavioral and dietary treatments for obesity have brought about a paradox in which most dietary programs demonstrate short-term weight loss, but the vast majority of people cannot maintain clinically significant weight loss over the long-term (13).

In this small group of GP patients, a power analysis indicated that all 11 subjects would have needed to complete the study to give it adequate power. Unfortunately, patients suffering from GP often develop poor eating habits in an attempt to control symptoms and are not compliant with diet change. An inherent weakness in the study lies in the fact that the subject population is very small. Another weakness is that the study relies in part on self-reported data, which is subject to self-serving bias. This error may have been compounded by the use of food databases to calculate the nutritional value of each food item.

However, among evaluable patients, some objective improvements were noted. Four of the evaluable five patients increased or decreased weight according to their goals, and the three patients who had developed GP as a result of diabetes reported decreased insulin requirements. Additionally, data from blood samples indicated a significant decrease in TNF-α and IL-8 levels upon implementation of the MD. This result is supported by various studies that have reported that polyunsaturated fatty acids exert anti-inflammatory effects and concluded that dietary supplementation with polyunsaturated fatty acids may inhibit the synthesis of IL-1 beta, IL-1 alpha, and TNF (14, 15).

The most noteworthy differences among evaluable subjects were observed in quality of life descriptions. While three of five participants reported that they disliked completing a food diary for reasons such as feeling self-conscious, judged, or watched, each of these patients also explained that maintaining food diary helped them become more aware of which foods they should supplement their diet with, as well as which foods tend to trigger GP symptoms. Certainly in the case of GP patients, increased cognizance of trigger foods can play a role in avoiding debilitating symptoms.

Decreases in TNF-α levels and some positive qualitative data collected from this small population, along with the consideration that gaining cognizance and changing old habits can enable acts of self-transformation, suggest that efforts to change poor established eating habits with the MD may be warranted. While some existing research does support the MD as an anti-inflammatory diet, its efficacy for GP patients requires further analysis and should take the role that sociocultural factors hold into consideration. For example, as today’s society increasingly applies nutrition to body image and body discipline as opposed to dietary disease, it is important to recognize that dieting holds human bodies to certain sometimes unrealistic standards, which is especially concerning in a population where food choices can already be so limited.

CONTACT SARAH BESSEN AT SARAH.Y.BESSEN.16@DARTMOUTH.EDU

References

7. Medtronic. 2003, Fact Sheet Enterra Therapy, Medtronic, Inc.
ARTICLE SUBMISSION FORM*
*Please scan and email this form with your research article to dujs@dartmouth.edu

Undergraduate Student:

Name: ___________________________  Graduation Year: _________________

School ___________________________  Department ___________________________

Research Article Title:

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Program which funded/supported the research ______________________________

I agree to give the Dartmouth Undergraduate Journal of Science the exclusive right to print this article:

Signature: ___________________________

Faculty Advisor:

Name: ___________________________  Department ___________________________

Please email dujs@dartmouth.edu comments on the quality of the research presented and the quality of the product, as well as if you endorse the student’s article for publication.

I permit this article to be published in the Dartmouth Undergraduate Journal of Science:

Signature: ___________________________

*The Dartmouth Undergraduate Journal of Science is copyrighted, and articles cannot be reproduced without the permission of the journal.

Visit our website at dujs.dartmouth.edu for more information
ARTICLE SUBMISSION

What are we looking for?
The DUJS is open to all types of submissions. We examine each article to see what it potentially contributes to the Journal and our goals. Our aim is to attract an audience diverse in both its scientific background and interest. To this end, articles generally fall into one of the following categories:

Research
This type of article parallels those found in professional journals. An abstract is expected in addition to clearly defined sections of problem statement, experiment, data analysis and concluding remarks. The intended audience can be expected to have interest and general knowledge of that particular discipline.

Review
A review article is typically geared towards a more general audience, and explores an area of scientific study (e.g. methods of cloning sheep, a summary of options for the Grand Unified Theory). It does not require any sort of personal experimentation by the author. A good example could be a research paper written for class.

Features (Reflection/Letter/Essay/Editorial)
Such an article may resemble a popular science article or an editorial, examining the interplay between science and society. These articles are aimed at a general audience and should include explanations of concepts that a basic science background may not provide.

Guidelines:
1. The length of the article should be under 3,000 words.
2. If it is a review or a research paper, the article must be validated by a member of the faculty. This statement can be sent via email to the DUJS account.
3. Any co-authors of the paper must approve of submission to the DUJS. It is your responsibility to contact the co-authors.
4. Any references and citations used must follow the Science Magazine format.
5. If you have chemical structures in your article, please take note of the American Chemical Society (ACS)’s specifications on the diagrams.

For more examples of these details and specifications, please see our website:
http://dujs.dartmouth.edu

For information on citing and references, please see:
http://dujs.dartmouth.edu/dujs-styleguide