Math and the mosquito

STATISTICS AND MACHINE LEARNING ARE THE NEW TOOLS TO DEAL WITH THIS SCOURGE OF THE TROPICS, WRITES **S ANANTHANARAYANAN**

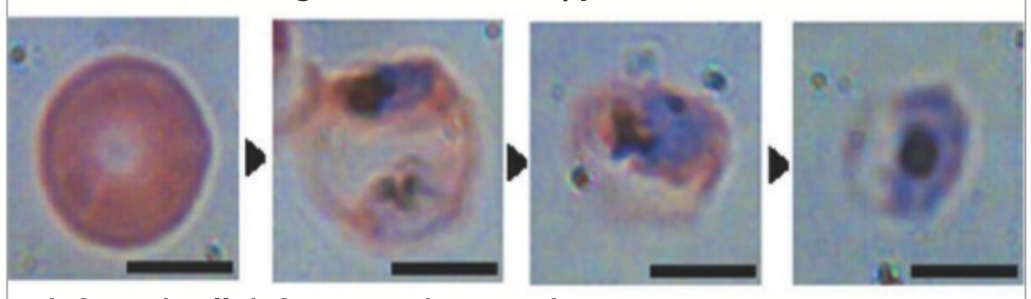
Most cases of malaria are curable provided the treatment is started early. The correct diagnosis, however, is only possible with a blood test and the current methods are time consuming and not available in many places. Malaria, hence, continues to rage in most of sub-Saharan Africa, Asia and Latin America, with over two million cases and some half a million deaths every year.

The report by Han Sang Park, Matthew T Rinehart, Katelyn A Walzer, Jen-Tsan Ashley Chi and Adam Wax of Duke University in the journal, *PLOS ONE*, of a fast and automated method of reliably detecting a major strain of the malaria parasite is, hence, of great significance. The method uses optical scanning of the blood sample and computerised, statistical analysis of an assortment of physical features of red blood corpuscles to estimate the presence of the parasite with reliability that compares well with the present method of manual analysis using a microscope.

Malaria is caused by a group of single-celled organisms in the Plasmodium group of species that enter the human body through a mosquito bite. The parasites move to the liver, where they reproduce in a form that then infects red blood corpuscles. Here again, they multiply till the blood cell bursts to release parasites to infect other blood cells. Some of the parasites develop into a form that can cause infection in other humans and these enter the body of a female mosquito when she bites an infected person. These grow in the mosquito's body and infect other humans, and so on (the male mosquito is inactive, as it does not feed on blood, only on plant nectar).

There are four major streams of Plasmodium, of which P. falciparum is the most prevalent and also the most deadly, one of its effects being on the brain, as cerebral malaria. In the standard procedure of examining the blood sample, a thick smear and a thin smear on a glass slide are stained and then checked under the microscope. The thick smear has a larger sample and an experienced technician can detect even traces of just five parasites per micro-litre of blood. But it is in the thin smear that the images are clearer, and the kind of parasite is more easily identified. It is important to know the kind of parasite present and both samples are generally needed. Picture 1 shows how the appearance changes with infection.

Picture 1 Bright field microscopy — From PLOS ONE



Uninfected cell deforms as the parasite matures

The problem is that the process takes time and there is the need for trained and skilled technicians. There are, therefore, many places where proper diagnosis is not practicable and treatment has to be based on a history of fever, as there are no specific symptoms. Treating for malaria by default can leave another actual ailment untreated and also lead to drug resistance, apart from loss of confidence in the health service.

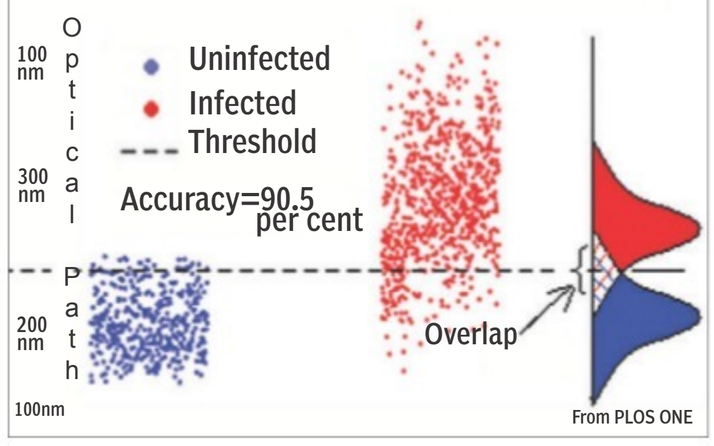
Automated method

The method developed by the Duke University team is to use a set of 23 visual/optical characteristics of infected red blood corpuscles by which an automated system could carry out rapid analysis of a sample. The features, such as size, ovality, its action on the path of light, the average and how frequent are deviations from the average of these features, symmetry, distribution of weight, are all found to be affected by infection. No one feature, however, is good enough to be relied upon to the extent necessary. It is not that this approach of analysing a brace of features has not been tried before, the paper says, but the earlier methods used bright field images after samples were fixed and stained.

The more recent method is do without staining but detect structure within transparent cell material by a technique called Quantitative Phase Imaging. In this method, part of a beam of light passes through the sample, while the other part takes an alternate route. The path through the cell, just microns thin, is longer to the order of the wavelength of light, which changes the phase of light waves that pass through the cell. This change can be detected and measured with the help of the interference of the two beams when they are brought together. The light beam is split into a widely separated spectrum, to start with, and the cell is analysed in this way in different wavelength ranges.

The Duke University team has made use of this method and also processed the mass of data that it produces with automated statistical analysis of the selected features of images of infected cells. Separate studies have shown that each of the 23 features do indicate the presence of infection, but diagnosis on this basis is not reliable. The value of the maximum optical path length, for instance, the paper says, is an indicator of whether the cell is infected, with 94.0 per cent specificity, 88.8 per cent sensitivity and 90.5 per cent accuracy.

Picture 2 MAXIMUM OPTICAL PATH LENGTH



From PLOS ONE

We can see in Picture 2 that there is an overlap that leads to inaccuracy if this feature alone is used to tell the two classes apart. Viewing and measuring the set of increasing numbers of infected cells by an automated system builds up data of the value of each feature, how closely the values are packed and the way they combine, in the case of infected cells. A machine learning system then uses established statistical methods to analyse large data to devise a method of telling the infected and normal populations apart with significantly higher accuracy.

In the normal course, analysis of data is carried out with the help of computers, based on a given formula, or an "algorithm". In machine learning, the system iterates through a succession of algorithms to see which one best fits the data available, and can hence be best relied upon to identify and classify fresh data. As the Duke University paper says, the system "builds a predictive model based on identified inputs as a teaching or learning set and classifies new datasets using a customised algorithm instead of following explicitly programmed instructions".

The authors of the paper discuss three statistical methods that could be used to enable the system to recognise infected cells. One is Linear Discriminator Classification, which applies different weightages to the values of identified features, and varies them to see what combination best tells the infected and normal cells apart. Another method is Logistic Regression, which is related to LDC but with different assumptions. And a third method is the Nearest Neighbour Classification, which is a method that tries to maximise a measure of similarity.

The authors carried out trials with real data samples and were able to show that LDC had a score of 98 per cent, 99 per cent and 99.7 per cent in accuracy of identification of infected cells at the different stages of infection that are shown in Picture 1, and this is higher than the success rate of other classification methods. Similarly, for sensitivity and specificity, the performance of LDC is generally superior.

The process could hence be used for quick and accurate detection of the stages of infection by P.falciparum, with a suitable device and mi nimal operator training, the paper says.

"This would permit rapid analysis of a blood sample at the point of care to assist the clinical decision of physicians," it says. This would be a crucial step for treatment and control of malaria in resource-limited areas where infection rates are the highest, a press release from Duke University says.

THE WRITER CAN BE CONTACTED AT response@simplescience.in

FAR MORE AMENABLE

RESTRICTION ENZYMES CLEAVE DNA MOLECULES AT SPECIFIC SITES, SAYS **TAPAN KUMAR MAITRA**

Since the hereditary similarities and differences observed among organisms derive from their DNA, one can expect the study of DNA molecules to yield important biological insights. Clues to a myriad mysteries — from control of gene expression within a cell to the evolution of new species — are to be found in the nucleotide sequences of genomic DNA. Most DNA molecules, however, are far too large to be studied intact. In fact, until the early 1970s, DNA was the most difficult biological molecule to analyse biochemically.

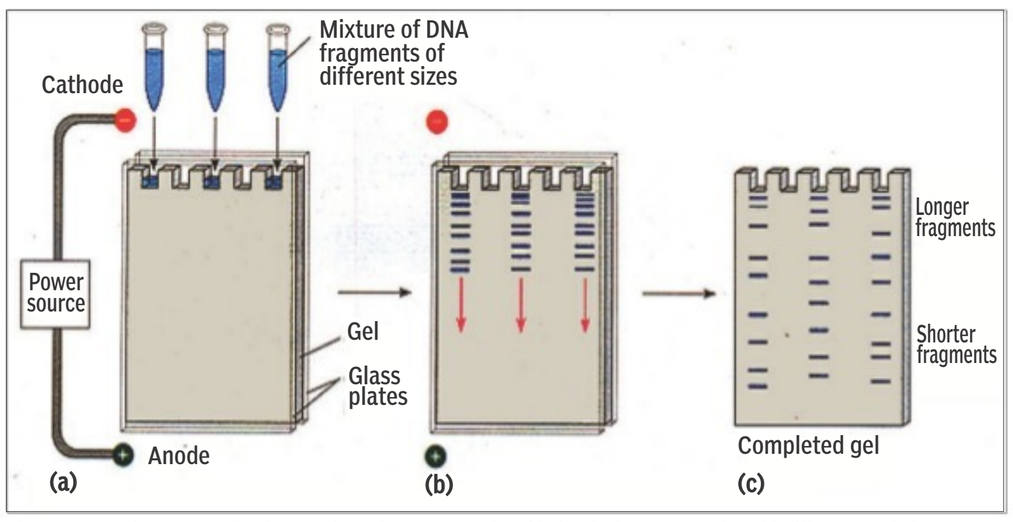
Eukaryotic DNA seemed especially intimidating, given the size of most eukaryotic genomes, and no method was known for cutting DNA at specific sites to yield reproducible fragments. The prospect of ever being able to identify, isolate, sequence or manipulate specific eukaryotic genes seemed unlikely. Yet in less than a decade, DNA became one of the easiest biological molecules with which to work.

This breakthrough was made possible by the discovery of restriction enzymes, which are isolated from bacteria that cut foreign DNA molecules at specific sites. The cutting action of a restriction enzyme generates a specific set of DNA pieces called *restriction fragments*. Each restriction enzyme cleaves double-stranded DNA only in places where it encounters a specific recognition sequence, called a restriction site, that is usually four or six (but may be eight or more) nucleotides long.

The frequency with which restriction sites occur in DNA is such that a given restriction enzyme will typically cleave DNA into fragments ranging from a few hundred to a few thousand base pairs in length. Fragments of these sizes are far more amenable to further manipulation than the enormously long DNA molecules from which they are generated.

Incubating a DNA sample with a specific restriction enzyme yields a collection of fragments of different sizes and to determine the number and lengths of such fragments and isolate individual fragments for further study one must be able to separate them from each other. The technique of choice for this purpose is gel electrophoresis, essentially the same method used for the separation of proteins and polypeptides. In fact, the procedure for DNA is even simpler than for proteins, because DNA molecules have an inherent negative charge (due to their phosphate groups) and therefore do not need to be treated with a negatively charged detergent to make them move toward the anode. Small DNA fragments are usually separated in *polyacrylamide* gels, whereas larger DNA fragments are separated in more porous gels made of *agarose*, a polysaccharide.

How does a researcher determine the order in which a set of restriction fragments is arranged in a DNA molecule? One approach involves treating the DNA with two or more restriction enzymes, alone and in combination, followed by gel electrophoresis to determine the size of the resulting DNA fragments. It would work for a simple DNA molecule cleaved with the restriction enzymes *EcoRl* and *Haelll*. In practice, restriction mapping usually involves data that are considerably more complex. In such situations, the DNA fragments produced by each restriction enzyme can be physically isolated — for example, by cutting the gel into slices and extracting the DNA from each slice. The isolated fragments are then individually cleaved with the second restriction enzyme, allowing the cleavage sites in each fragment to be analysed separately.



The three test tubes (a) contain mixtures of DNA fragments produced by incubating DNA samples with different restriction enzymes. To fractionate a DNA preparation containing fragments of various sizes, a small sample of the preparation is applied to the top of a gel. An electrical potential of several hundred volts is then applied across the gel, such that the anode (the positive electrode) is at the bottom of the gel and the cathode (the negative electrode) is at the top. The DNA fragments (b) in the applied sample migrate toward the anode, with shorter fragments migrating more rapidly than larger ones. After enough time is allowed for separating fragments (c), the gel is removed and stained with a dye such as ethidium bromide, which binds to the DNA fragments and causes them to fluoresce under ultraviolet light. Alternatively, autoradiography can be used to locate the DNA bands in the gel, provided the DNA is radioactively labelled.

The frequency with which restriction sites occur in DNA is such that a given restriction enzyme will typically cleave DNA into fragments ranging from a few hundred to a few thousand base pairs in length. Fragments of these sizes are far more amenable to further manipulation than the enormously long DNA molecules from which they are generated.

Incubating a DNA sample with a specific restric-

THE WRITER IS ASSOCIATE PROFESSOR, HEAD, DEPARTMENT OF BOTANY, ANANDA MOHAN COLLEGE, KOLKATA, AND ALSO FELLOW, BOTANICAL SOCIETY OF BENGAL, HE CAN BE CONTACTED AT tapan-maitra59@yahoo.co.in

Brain implant does it

IAN JOHNSTON REPORTS ON A PARALYSED MAN BEING ABLE TO FEEL THROUGH ROBOTIC FINGERS IN A WORLD-FIRST BREAKTHROUGH

A 28-year-old man left paralysed after a car accident has been able to feel as though he was touching something with his fingers after a robotic arm was connected directly to his brain in a world-first breakthrough. Nathan Copeland, who was injured after crashing his car on a rainy night in Pennsylvania when he was 18, spoke of experiencing a "really weird sensation" as he touched things. He said it felt like "my fingers" were being touched or pushed.

Copeland is able to feel using the robotic arm because it is connected to microelectrodes about half the size of a shirt button that were surgically implanted in his brain. Before the operation, imaging techniques were used to identify the exact places that corresponded to feelings in his fingers and palm.

The discovery that people can regain some sensations using a so-called "computer-brain interface" could revolutionise the treatment of paralysis.

Earlier this year, the Walk Again Project in Brazil discovered that people left paralysed by severe spinal cord injuries could recover the ability to move their legs after training in an exoskeleton linked to their brain. That project was designed to enable people to walk by controlling the exoskeleton with their minds, but one of the subjects was able to walk again using crutches.

Professor Robert Gaunt of Pittsburgh University, who led the team that treated Copeland, said they were trying to make use of the brain's natural abilities."The ultimate goal is to create a system which moves and feels just like a natural arm would," he said. "We have a long way to go to get there, but this is a great start."

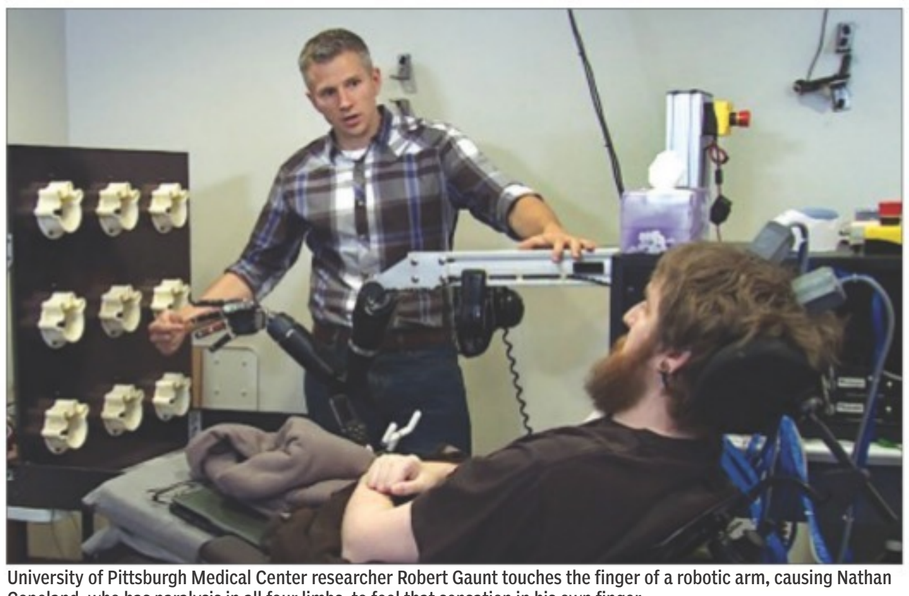
His colleague, Professor Andrew Schwatz, said the most important finding was that the system could create a "natural sensation". But he added, "There is still a lot of research that needs to be carried out to better understand the stimulation patterns needed to help patients make better movements."

When his accident happened, Copeland was in his first year of college studying for a degree in nanofabrication. He tried to continue his studies, but health problems forced him to put them on hold. One of the first things he did after he was injured was to enrol on the Pitt School of Medicine's registry of patients willing to participate in clinical trials. Ten years later, that led him to have the operation to fit the implants in his brain and rediscover what it is like to reach out and touch things. "I can feel just about every finger — it's a really weird sensation," he said, speaking a month after the operation. "Sometimes it feels electrical and sometimes its pressure, but for the most part, I can tell most of the fingers with definite precision. It feels like my fingers are getting touched or pushed."

In a video interview, Copeland, who can move his upper arms but has no sensation or movement in his lower arms and hands, added, "I usually feel it in the base of my fingers, in my finger pads, usually a tingle or some pressure like someone was squeezing. A couple of electrodes feel like they are on my knuckles... there are a couple of electrodes that feel like a regular touch. It's never been painful. It's just kind of a tingle, it's not really pleasant or unpleasant."

The research was described in a paper published in the journal *Science Translational Medicine*.

THE INDEPENDENT

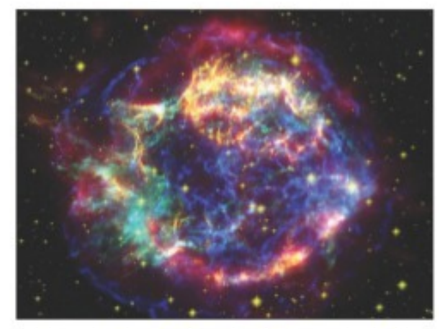


University of Pittsburgh Medical Center researcher Robert Gaunt touches the finger of a robotic arm, causing Nathan Copeland, who has paralysis in all four limbs, to feel that sensation in his own finger..

PLUS POINTS

New space nation

Scientists have launched Asgardia, the first new space nation, and anyone can become a citizen of it. What it hopes to do is save the human race twice over. First, it will look to protect it from warfare in space; and, second, it will try



A false colour image of Cassiopeia A comprised with data from the Spitzer and Hubble Space Telescopes and the Chandra X-Ray observatory.

and keep humanity safe from the dangers coming from outside, protecting us from threats like space debris and asteroid collisions.

Those behind Asgardia hope that creating the country is the first step of a new era in the space age. And they intend to start that by sending up rockets. The country will send its first satellite into space in 2017. From there, Agardia hopes to "open up access to space for commerce, science and peoples of all countries on earth". The scientists behind the plan launched it in Paris last week and named it after the city that was ruled by Odin from Valhalla in Norse mythology.

It is being led by Igor Ashurbeyli, who leads the Aerospace International Research Center in Vienna and is chairman of Unesco's science of space committee. Asgardia was created in consultation with "globally renowned scientists, engineers, entrepreneurs and legal experts", according to those behind it, and it is hoping to become a fully recognised country, said Ashburbeyli. When it does so, it will be able to promote values central to scientists, he claimed. "Any human living on earth can become a citizen of Asgardia," the site's citizenship page reads. It includes a link to a form to join the country, which is done just by handing over your first name, last name, email and country.

ANDREW GRIFFIN/THE INDEPENDENT

Royal mystery solved

In 1934, King Albert I of Belgium died in a rock-climbing accident at the age of 58. Eighty years later, Belgian television journalist Reinout Goddyn purchased some bloodstained leaves supposedly collected from the site of the accident because he wanted to settle the conspiracy theories that circulated after the king's death, which no one had witnessed.

In 2014, Goddyn reached out to Dieter Deforce, director of the Laboratory of Pharmaceutical Biotechnology, Forensic DNA, at the University of Ghent. "He asked if I could do a DNA analysis to prove if it was the blood of the king or not," recalls Deforce, who was hesitant to help because doing so would require a sample from one of the king's relatives. But he told Goddyn he could test the bloody leaves to see if the DNA was human or not. Found to be human, the two appeared together on a Flemish television programme to discuss the finding later that year.

Watching that broadcast was forensic geneticist Maarten Larmuseau, a postdoc at the University of Leuven, Belgium, and he was not satisfied. "That was not the end of the story," he said. "You had to identify with DNA if the blood indeed belonged to the king or not." He reached out to Deforce, who connected him with Goddyn, who immediately agreed to the project. Larmuseau also contacted two distant relatives of King Albert I — one on the father's side and one on his mother's side. Because the blood was 80 years old, the DNA was highly degraded, but thanks to advances in technologies to deal with such poor quality samples, Larmuseau and his colleagues were able to run the comparison on 42 Y-chromosome loci and sequence the entire mitochondrial genome. The DNA from the blood was a match with both the Y-chromosome loci of his paternal relative and the mitochondrial sequence of his maternal relative. Once the researchers had their results, "that's when the problems started". The raw data contained the Belgian royal family's sensitive genetic information. Comparisons of the Y-chromosome data with other family members' DNA sequences, for example, could identify cases of infidelity, said Larmuseau.

Larmuseau said that, in addition to personal and historical curiosity, the ethical issues surrounding the publication of historical genetic data was what motivated him to reach out to Deforce and Goddyn in the first place.

JEFF AKST/THE SCIENTIST