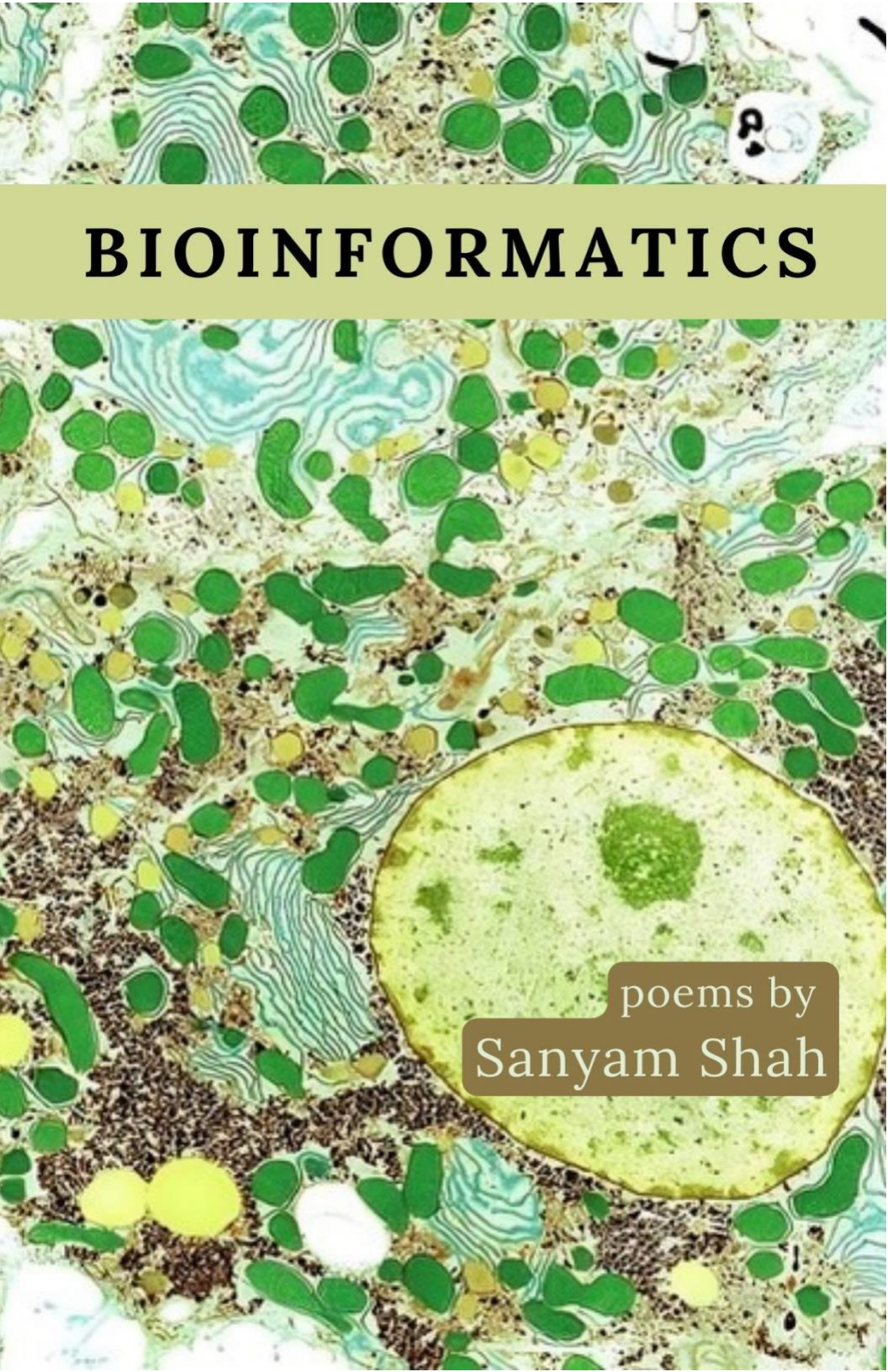
BIOINFORMATICS



poems by

Sanyam Shah

© 2025 Sanyam Shah.

This collection is meant to be shared. You may freely distribute, quote, or reproduce these poems for noncommercial purposes — just please credit the author.

ISBN: 978-93-343-4923-8

This collection is registered with the United States Copyright Office.

First Edition – 2025

Self-Published

Introduction

Purpose and Inspiration

Over the past two years, I've been amazed by the beauty and interconnectedness of the natural world, and this appreciation has only grown as I've studied bioinformatics. Growing up, I had so many interests—math, biology, programming, AI, statistics—that I struggled to pick just one thing to pursue. For a long time, I thought choosing a career meant giving up on all the other things I loved. Then, I stumbled across bioinformatics, and everything clicked. It was the perfect way to combine all my passions into one field.

As I studied bioinformatics, I realized that it has so much potential to change the world, but it's also not always easy for people outside the field to understand. Science can feel intimidating because of all the technical terms and complicated ideas, which can make it hard for people to connect with. That's why I chose to focus my honors thesis on using poetry to share the beauty and creativity of bioinformatics. I wanted to bridge the gap between science and the humanities/poetry, which is arguably a more universally understood and inspiring language.

Methodology

When writing this poetry collection, I followed a simple process. First, I picked a specific area of bioinformatics to focus on. If it was a topic that I

wasn't familiar with, I did some research to learn more about it. I'd read papers, explore studies, and take notes on any ideas, metaphors, or imagery that came to mind.

Once I had a rough understanding, I'd sit down and start writing. Most of the time, I didn't overthink it—I just let the ideas flow and made adjustments as I went. I'd add, remove, or change things until I had a rough draft. This process usually took about 30 minutes per poem. After writing a few poems, I'd send them to Dr Rosemarie Dombrowski for feedback, and her suggestions helped me refine my work and think differently about future poems.

Outcomes

The result of this process is a collection of 15 poems, each about a different area of bioinformatics. These poems explore the kinds of research that can improve human health and push science forward. My main goal was to make science less intimidating and more accessible. I wanted to break down the barrier that big words and technical language often create, which can scare people away from learning more about science.

By blending art and science, I hope these poems help people see the beauty behind what we do in bioinformatics. Science isn't just about facts and figures—it's about creativity, problem-solving, and discovering the unknown. My poems aim to capture that and make it relatable for anyone, no matter their background.

Future Applicability

To the best of my knowledge, this is the first poetry collection focused on bioinformatics. But the idea of using creative writing to explain science could be applied to other fields too. It's a way to help people understand research and its importance, even if they don't have a scientific background.

I hope this work inspires more people to see the art in science and maybe even pursue research themselves. It could also help bridge the gap between scientists and the public, showing that science isn't just for a select few but something we can all appreciate and connect with. By making even the most complex topics more accessible, we might inspire curiosity, creativity, and collaboration in ways we haven't yet imagined.

Sequence Alignment I

Sequence alignment is a way of arranging the sequences of DNA, RNA, or protein to identify regions of similarity.

We align the sequences of life,
hoping to fill in the gaps
that define our history and identity.

We look for patterns
in caves of information.

We look for differences—
a mutation that separates
life from death,
cancer from no cancer,
joy from misery—
so that someday,
when someone's sister
cries in a hospital room,
I can point to a sequence
and say *Here, here's the spot*
where we'll make it right.

RNA Seq Analysis

RNA sequencing (RNA-seq) analysis is a process that uses specialized computational tools to analyze the nucleotide sequence of RNA molecules and their quantities.

It echoes through the body
like sound in a cave,
the rhythm of life,
instructions on how to survive.

I'm amazed when I look
at the computer, each sequence
telling a story of life and death,

I wake up at midnight, unable to sleep
I sit at my laptop and analyze a sequence—
Try to solve the maze, find clues
that might have revealed my friend's disease,
hinted at the cure.

Comparative Genomics

A field that compares the DNA sequences of multiple species to identify similarities and differences.

We lay genomes side by side,
compare the ancient texts of our existence,
diluted and perfected by time.

I observe similarities
in the DNA of the mouse—
genes that make our hearts beat,
eyes blink, lungs breathe.

The secrets lie in the one-percent,
the genetic difference that allows me
to write poetry, worry about
my grandma's health and make her tea.

I am amazed when I see a sequence
that tells the story of a billion years,
providing clues to the first cell that made us.

I am hopeful that one day,
a difference will show itself,
perhaps in a bat's code,
explaining the absence of cancer,
a pattern that shields us from pain.

Metagenomics

A method for studying the structure and function of microbial communities by analyzing the DNA extracted from environmental samples

Small but not insignificant,
microbes roam aimlessly,
sailing the winds on invisible currents,
in search of vulnerable shores.

They enter the body,
coursing the stream of blood,
hijacking new cells, injecting its DNA.
A poor worker, using its energy
to live, to reproduce.

It flourishes while a young woman cries
and a grandma worries,
placing wet towels on her forehead.

We sample the environment,
that harbors these crews.
By sequencing DNA
from the earth and the skies,
we identify our foes.

But the result is in a never-ending war—
we fight, knowing that at the end of the day,
we can stop a loved one's cries
as the host recovers
and the
wet towels disappear.

Systems Biology I

The computational and mathematical analysis of Biological Systems.

Equations: so small,
so simple,
encompass nature
and all of its rules.

If we change the number
in equations,
an ecosystem dies,
a virus survives,
a species evolves.

In other words,
the probability
of your loved one's death
can be represented
by an equation.

In other words,
you would not be here
to experience your joys
and your sorrows if not for
the correct combination of numbers.

Some evenings, I realize
that if we knew how to simulate disease,
my grandpa would still be here,
making me tea to brighten
the sky within me.

It is our job to contain a forest in a jar,
to travel through the
infinite mess of equations,
to find the key to the universe.

Sequence Alignment II

Sequence alignment is a way of arranging the sequences of DNA, RNA, or protein to identify regions of similarity.

There are gaps in our knowledge,
in our history and that of our fellow aliens,
missing branches in the tree of life.

We compare the DNA of species
to find where they match
and to understand how close they are.

I like to think of it as comparing ingredients
of dishes. The more common
the ingredients, the more common the taste.

As we do this, the branches grow,
leaves emerge, and the tree is complete,
just like our identities.

Metabolomics I

The study of interactions between small molecules in the body.

Within you are molecules,
millions interacting
in perfect harmony,
like each pixel
working together
to create a perfect image.

The molecules divide,
join and multiply,
every second a new state,
a kaleidoscope spinning
new patterns.

It surprises me
that a billion cells,
a trillion molecules
work effortlessly to make
a singular being.

What I'm saying is that
chaos keeps us alive.

Proteomics

The study of proteins in a cell, tissue or an organ.

Proteins are complex puzzles
solved by nature,

A slight distortion in that perfection,
a small misfolding of the art
and your family screams,
tears flowing from the blue eyes
of your beautiful child
as they try desperately
to make you remember them.

We modify shapes, curious about
how this change can evolve us,
like trying to create the best robot out of Legos.

We desperately look for a shape,
a protein that can cure
someone's grandma, sister or brother.

A protein that stops tumors,
another one that kills viruses,
an adventure to find the geometry
of a protein that saves us from pain.

DNA

The molecule that contains the genetic information necessary for an organism's function.

DNA is a river of data
flowing through generations.
Each droplet a sequence.
Take a bucketful
and you uncover the secret
of sight, the ability to witness
the art and complexity
of creation.

Dive deeper.
Take a swim.
Let your skin absorb the water.
You will be lost and confused.
You try to comprehend your existence.
You try to rise to the surface,
but the closer you get,
the farther the surface seems.

Phylogenetics I

The study of evolutionary relationships between species using DNA sequencing.

Humans are like leaves on a tree—
not knowing the branches or the roots
through which we ascend.
When the wind blows,
we separate from life
so easily, so peacefully.

When the sun rises, the leaves grow.
They look up and see the future.
They are unaware of the grip
of the past that holds them in place.

The leaves shine
and the roots rest in the darkness.
When the sun shines,
the rays escape.
They light up the path
as we dig for the roots.

Phylogenetics II

The study of evolutionary relationships between species using DNA sequencing.

We study a thread
that connects us—
interwoven, beautifully ordered,
created with love and attention
like my grandmother's blue sweater.

We travel across the thread endlessly.
In discovering stories, we discover ourselves.

Our genes are the basis of our survival.
There is no certainty, no clarity,
just a compass with no needle.

Metabolomics II

The study of interactions between small molecules in the body.

Reactions, small yet important,
occur infinitely to keep you alive.

Cells, molecules, tissues
all work in chaos
to keep you at peace.

When I feel low,
I think about the drama going on inside,
keeping me alive, and my problems
become as small as those cells.

We study these molecules
who I will never know—
my loyal crew,
the managers of all things molecular.
I owe it to them to be happy.

Transcriptomics

The study of RNA molecules in the cell. RNA is formed by transcription of DNA. It is responsible for the coding of all the proteins in the body, regulating all the genes, etc.

Let's go on a journey:
You are a small molecule.
Inside the cell,
you see a twisted thread,
so tightly packed
that if stretched out,
it would be two meters long.

Every cell in your body
has this thread.
You move closer,
and as you travel,
you read the code
one piece at a time.
Walking along the thread,
you stop to look:
here, a code that colors
your eyes blue.
There, a code that helps
digest fat.
Tired from the journey,
you pause.
You rest at a spot
that gives instructions
on how to build proteins.

Ahead, the thread
keeps going endlessly,
doing many things
at once.

You step back
so you can see it all—
everything working together
in perfect harmony,
the amazing design
of the body that holds you.

RNA Seq Analysis II

RNA sequencing (RNA-seq) analysis is a process that uses specialized computational tools to analyze the nucleotide sequence of RNA molecules and their quantities.

Small threads,
invisible to the human eye,
provide the light.

These threads contain codes:
for proteins,
for the love you feel
for your mother,
every emotion,
physical trait, and response.
Four letters telling a story
of your past, present, and future.

We study the sequence to determine
how change affects us.

Wanting to learn more, we sleep hungry,
knowing the amount of knowledge,
yet to be found.
We are desperate for more,
hungry for the knowledge
yet to be found.

How unaware we are
of our own compulsions.
How much free will,
if any, is allowed?

How often does the mind win
in this infinite war against our genes?

Systems Biology II

The computational and mathematical analysis of Biological Systems.

Machines
cold and dead
understand the world
better than the living.

We put the lives
of our mothers
and sisters in their hands
hoping that they are right
(Because we know
We never will be).

We experiment on creatures
we think are below us—
On rats, bats,
even other humans
desperate to understand what we are
and why we exist.

I hope the machines
remain cold and dead.
I fear that if a pinch of humanity
becomes part of them,
we might become the rats, the bats.

About the Author

Sanyam Shah is a bioinformatics researcher passionate about the intersection of biology, mathematics, and technology. He has explored various fields within bioinformatics, including clinical informatics, metagenomics, and machine learning applications. As an active researcher, Sanyam has authored multiple peer-reviewed publications.

Alongside research, Sanyam is deeply interested in storytelling, poetry, psychology, and mental health. In addition to this poetry collection inspired by bioinformatics, he is also the author of a psychological fiction novel exploring depression, loneliness, identity, connection, and suicidal ideation. Through both scientific and creative work, he hopes to better understand and support people; not only biologically, but emotionally.

Dedicated to advancing genomics and making impactful discoveries, Sanyam aspires to pursue a career in bioinformatics research to continue exploring the vast potential of bioinformatics to improve human health.

Email him at positivesanyam@gmail.com