SelfDecode™

# Easy to use, affordable, and on-demand AI doctor.

# Thesis

Empower consumers, businesses, researchers, and physicians to utilize the full power of genomics through an **AI-powered precision medicine platform**.

# Vision

Make precision medicine a routine, affordable and actionable part of healthcare to **prevent diseases and radically extend healthspan**.

# Problem

We've barely scratched the surface in uncovering the insights hidden in our genes.

## Genomic Data is Massive & Complex

By 2025, researchers estimate there will be 2 billion genomes sequenced equaling roughly 40 exabytes of data*. This data must be stored, transmitted and processed at a cost well over $1 trillion / year.

## Tools are not Available for Accurate Risk Scores

Due to an inability to handle or analyze the data properly, Polygenic Risk Scores (PRS) are not powerful enough for clinical or pharmaceutical applications. A PRS also needs to combine with environmental and lab data to get a much more accurate risk assessment. This aggravates the data complexity.

*Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, et al. (2015) Big Data: Astronomical or Genomical? PLoS Biol 13(7): e1002195. https://doi.org/10.1371/journal.pbio.1002195 (2 billion whole genomes totaling 40 exabytes stored on cloud at 2.1 cents per GB plus processing costs)

## Core Problem

It's not financially or technologically feasible for individual business entities or genomics researchers to store and process genomic data - even big pharmaceutical companies (as well as hospitals, academic institutions, health and life insurance companies). This situation stifles genomics research and pharmaceutical development and prevents precision medicine from widespread application.

# Solution

Develop an end-to-end platform for health apps, biotech, pharmaceutical and insurance companies, hospitals and academic institutions to **analyze, store, and process** genomic and phenotypic data.

STRATEGY

1

Focus on the DTC business which generates cash flow while accumulating an extensive genomics and phenotypic data set. Use customer data to refine the PRS.

2

Continue to build out tools for practitioners and users that will also have utility for B2B customers (clinical decision support system & increase member engagement with health insurance companies).

3

Develop a platform for population genomics with utility for the DTC business and strong market potential for the precision medicine industry, including data compression, genetic imputation, improved ancestry, genome simulation and dramatically better Polygenic Risk Scores.

TRACTION

# SelfDecode DTC

$2.5M trailing 12 month revenues (ending July 2021)

Est. $400 Lifetime Value of a Customer (LTV).

15,000 current subscribers

**Recurring Revenue:**

Predictable revenue stream from subscription sales model.

**High Margins & Scalable:**

High Avg Order Value of $254 - Low Cost of Goods Sold of ~$20 = $234 Gross Profit/Customer.

**Highly Rated:**

4.6 Trustpilot score, 98% retention rate, 75 Net Promoter Score (as of July 2021)

**Operational Excellence:**

Bootstrapped to multimillion dollar revenue with only small seed investments (totaling $450K).

TRACTION

# SelfDecode's R&D Team

One of the best-qualified science teams in the genomics industry. 36 scientists, MDs, PhD, AI and Big Data engineers.

We've built an industry-leading genomics tool set. We plan to file 5 patents within 6 months and publish our findings in leading peer-reviewed journals.

## Polygenic Risk Score (PRS):

We have a prototype PRS tool significantly more accurate than current gold standard algorithms. In addition to being more accurate, our PRS tool calculates individual PRS at 1/80th the cost of conventional algorithms.

## Genomic & Phenotypic Simulation:

Solves the Data Privacy problem by allowing us to model millions of genomes in the public data set without having to pay for it. The modeled data can be sold to companies, in addition to improving our PRS.

## Data Compression:

Our genomics optimized data compression tool currently compresses at a ratio of 1/5000, 10X better than the current industry record.

## Data Engineering:

Normal databases can only hold 20,000 rows for real time data queries. Our data engineering tools enable us to store and query against data records comprising 80M SNPs.

## Imputation:

Our imputation tool does rare variants better than any other imputation currently available. Rare variants are the most relevant drug targets. Our imputation cost is currently 1/5 the industry norm.

## Ancestry:

Our ancestry tool is based on the multi-ethnic, multi-generational SelfDecode dataset, enabling us to do better multi-ethnic analysis. This is required for FDA approvals of diagnostics tests.

# Why Now

## The era of precision health has begun. This shift is made possible by;

- Availability of government funded massive data sets (All of Us, UK Biobank, etc...)
- The cost of training AI models has dropped dramatically.
- Cloud computing revolution makes it feasible to develop AI models using cloud resources on an as-need basis.
- Increased consumer/patient awareness of DNA testing and its medical relevance.
- Health plans, hospitals and clinicians are incorporating genetic data into their business models.

## We are uniquely positioned to take advantage of this opportunity because;

- We have a full R&D team comprised of genius level genomics scientists, bioinformaticians, AI scientists and big data experts.
- From our position as a leading DTC genetic testing service we have a large multi-ethnic, multi-generational data set for our science team to work with.

# Market

GMI Insights reported that the digital health market is expected to grow to over $639 billion by 2026.

Accenture 2020 Digital Health Consumer Survey showed that 62% of consumers are open to using digital health services.

Preventive healthcare technologies and services market will be worth $432 billion by 2024.

**SelfDecode would like to not only service some of this market share directly, but even more so leverage our technology to provide the capability for other companies to provide personalized health services to their customers.**

# Competition

## Most DTC Genetic Testing Services Are Both Inaccurate and Unhelpful:

DTC testing services fail to account for phenotypic data, lab test results, or lifestyle factors. They also don't provide personalized recommendations.

DTC genetic testing services use just a few alleles to drive their conclusions. New studies show that millions of variants can affect a given trait.

All DTC genetic testing services (except 23andme) do not have science and engineering teams or infrastructure capable of implementing the latest genetic analysis.

COMPETITION

| DTC Competitors | Target market | Strengths | Weaknesses |
|---|---|---|---|
| **23andme** | Health & Ancestry | First mover status, large data set, good designs, relatively good science | **No recommendations or personalization based on genes, no lab tests. Science is limited in scope for drug targets.** |
| **Circle** | Health & Ancestry | Good UI/UX | **No legitimate science** |
| **Ancestry** | Ancestry | Good Ancestry product | **No health product** |
| **Human Longevity** | Health | Comprehensive testing | **Unaffordable, no focus on recommendations or analysis** |
| **Nebula** | Health | Whole genome for cheap | **Very bad science, no recommendations, poor UI/UX** |

COMPETITION

| Genomics Competitors | Target market | Strengths | Weaknesses |
|---|---|---|---|
| **23andme** | Pharma | Well funded, large science team, FDA authorizations | **Narrowly focused on drug discovery for GSK (exclusive agreement).** |
| **Allelica** | PRS | Focused effort, partnered with Merck | **Small team, limited capability, weak in the business and marketing** |
| **Color** | End-to-end study support, health insurance, employers | Big biz dev team | **Small actual science team, really a service company, potential customer** |
| **Helix** | Population genomics studies, health plans | Large dataset, tools platform, partnered with Illumina, trials with healthcare providers | **Service company, expensive cost structure, potential customer** |
| **Public domain tools from academia.** | Research | Free | **Limited support, poor performance, no integration, restrictions on commercial application** |

# Financing

Cash on hand:

**$6.5 million**

Funding:

**$7.5 million**
(expecting to receive another 500k yr end from Wefunder)

Additional Funding:

**$2-3 million**
to close this round of funding.

# Use of funds:

1. Build out DTC Product & Marketing team
2. Build Out Genomics Platform to meet business needs
3. Conduct health economic and validation studies
4. Hire B2B sales team

# Team

**88 Total Team Members (80 full time, 8 part time):**
including 30 scientists (PhDs, Masters) and 38 software architects and engineers.

## Joseph Cohen

FOUNDER, CEO

8+ years of professional executive experience in the health and biotech industry with a proven track record of success. Expertise in operations, HR, management, product, marketing, finance & legal. Internationally recognized speaker and biohacker.

Growing up, I suffered from inflammation, brain fog, fatigue, digestive problems, anxiety, depression, & other issues that caused me to rack up medical bills. Finally, I started studying my genes and found solutions that actually worked to get rid of my health issues. I knew if personalized healthcare could work for me, it could work for others.

## Dr. Puya Yazdi, MD

CHIEF SCIENTIFIC OFFICER

MD from USC, Stanford Residency. 15+ years of success in R&D, IP acquisition & science publications in genomics, bioinformatics, biotech & precision medicine. 10+ precision medicine products. 7+ years of executive experience.

## Ralph Kenney

CHIEF OPERATING OFFICER

20+ years of successful C-level experience in technology, healthcare, and e-commerce (including AT&T). As COO, he has taken a previous company from start-up to successful IPO in 6 years. Founded/CEO of a profitable 7 figure business.

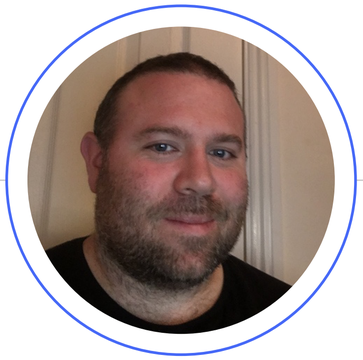
# Team

**88 Total Team Members (80 full time, 8 part time):**
including 30 scientists (PhDs, Masters) and 38 software architects and engineers.

## Manfred Grabherr,  PhD

VP OF BIOINFORMATICS AND AI

PhD & Prof at top Swedish uni, w/ 30+ yrs of experience in Genomics, AI & Bioinformatics. Worked at MIT & Harvard's Broad Institute, 60 publications in top journals (Nature, Science, etc), 35k citations & numerous patents. Had prior successful exits.



## Charles Manson, PhD

AI/MACHINE LEARNING ENGINEER & STATISTICIAN

PhD in Math. Attended Cambridge, NYU & Warwick Universities. Previous AI/Deep learning startup founder. A top student of top mathematician - Martin Hairer (Fields Medalist, $3M Breakthrough Math Prize).

SelfDecode™

INVESTOR PITCH DECK

OCT 2021