I have made significant scientific contributions in multiple distinct fields. I am an author of 13 peer-reviewed publications, six of which I am the lead, corresponding or senior author. These works include several foundational and innovative studies important within their respective fields.

My doctoral research with Dr. Dixie Mager on the role of endogenous retroviruses (ERVs) in human cancer **established the concept of *"onco-exaptation"***. This is the concept that genetic and epigenetic evolutionary pressures shape how ERVs and other transposable elements can be co-opted into novel transcriptional units during cancer development. The seminal *Oncogene* paper coining this term, and the *Mobile DNA* review maturing this concept have been cited 39 and 78 times, respectively. The influence of my work on the direction of thought within the field of ERVs in cancer is demonstrated by these citations and the application of the "onco-exaptation" concept by other research groups.

As the world grinded to a halt in the wake of the COVID-19 pandemic, I saw it as the responsibility of all scientists to fight this scourge. In this light, **I began *Serratus*, a computational project to analyze all public high-throughput sequencing data to uncover coronavirus sequences**. I personally secured $197,000 CAD of research funding am leading an international team of 13 scientists to develop the world's foremost sequence-alignment computing architecture. With *Serratus* we deployed a cloud-backed 22,000 CPU cluster capable of aligning >1 million sequencing datasets *per day*. This amounted to 713 CPU-years of computing in a mere 7 days, aligning an unprecedented 3.84 million sequencing datasets. The novel coronaviruses we discovered offer critical insight into the evolutionary origins and zoonotic dynamics of this viral family. One newly discovered group of corona-like viruses contained segmented genomes, an unexpected finding which challenges the textbook definition of a coronavirus. These segemented corona-like viruses may undergo reassortment (like influenza) and thus are highly prone to cross-species transmission, changing our textbook understanding of coronavirus biology. The manuscript of this work is currently in review at a leading journal and I am the senior author.

My most significant scientific contribution to date is in the field of ribosomal RNA genetics. I independently initiated research on human ribosomal DNA (rDNA) and found it is variable both within individuals, and across human populations. Such variation directly contravened the dogmatic notion in the field stating rDNA copies are homogenous (via gene conversion) and variation of this ancient gene is not tolerated. To explore the functional consequences of rDNA variation I turned to cancer biology, hypothesizing that oncogenic variants would be selected for in course of cancer evolution. Within this system I made the very surprising discovery that the 1.5 billion year conserved rRNA modification, 1-methyl-3-(α-amino-α-carboxyl-propyl) pseudouridine, is sub-stoichiometrically lost in colorectal cancers (CRC). This is noteworthy for two reasons; first, previous literature stated that this RNA hyper-modification was invariable in all eukaryotic ribosomes, which is not true; second, a staggering <u>45% of CRC patients</u> show sub-physiological modification, meaning this molecular variation is penetrant to an extent comparable only to the most severe driver mutations, such as those in *TP53*, *KRAS* and *APC.* I am the lead/corresponding author on the *Cell Reports* publication reporting these findings.

In this proposal I will continue and expand this exciting research on rRNA variation at the University of Cambridge, joining Dr. Warren's vibrant laboratory offering synergistic expertise on ribosome biology. Together we will explore the molecular and structural underpinnings of these findings to understand their consequences on cancer biology and its potential as a chemotherapeutic target.