Activities Planned Under This Award:

Percentage of time to be spent under each category:

	Year 1	**Year 2**	**Year 3**
Research	95	95	95
Others	5	5	5

The majority of my time will be dedicated to the proposed research project during the funding period of this fellowship. However, I will also engage in the following activities to supplement my training:

All course work and teaching requirements of my PhD program have been fulfilled during my first two years of graduate study. However, I will continue to enroll in the responsible conduct of research training curriculum at Harvard Medical School (HMS).

To stay informed about scientific progress in bioinformatics and biomedicine and to develop my communication and presentation skills, I will be attending a variety of seminars, workshops, and retreats hosted by HMS and other affiliated institutions. My PhD program organizes a work-in-process seminar series in which students from different laboratories can meet and share their research on a regular basis. Regular journal clubs and faculty talks are also available at HMS. I will travel to local and national conferences related to my research, including the ASHG Annual Meeting and the Single Cell Analysis Investigators Meeting.

Timeline of research:

Year 1:	**Year 2:**	**Year 3:**
- Gather and process datasets - Code hierarchical Bayesian models for CNV and SNV inference (Aim 1) - Benchmark model sensitivity and precision using simulations and artificial mixtures of K562 and GM12878 cells (Aim 1) - Code Bayesian model phylogenetic reconstruction model - Benchmark model against regular hierarchical clustering using simulations and artificial mixtures of K562 and GM12878 cells (Aim 2)	- Apply hierarchical Bayesian models for CNV and SNV inference to CLL datasets (Aim 2) - Bayesian model phylogenetic reconstruction model to CLL datasets (Aim 2) - Differential expression and gene set enrichment analysis on putative CLL subpopulations (Aim 3) - Compare genes and gene sets across patient and across time points (Aim 3)	- Functional validation of differential gene expression using targeted rt-qPCR and other methods at the discretion of collaborators (Aim 3) - Finish analysis - Package up code - Create website and supporting features - Prepare manuscript - Submit manuscript - Submit software to BioConductor and other code repositories - Perform additional analysis, revise manuscript, and resubmit as needed