Markov Chain Clustering is a powerful statistical approach that has been widely applied in the field of bioinformatics. Stijn Van Dongen invented the algorithm at the Center of Mathematics and Computer Science in Netherlands. Mentzen et all, have studied the organization of expression of Arabidopsis genome using MCL and have asserted the importance of the study. Their data set had 22,746 genes which, when pre-processed for clustering will result in a matrix 2 Gigabytes in size on computer memory. Irrespective of the time taken to build this matrix, the time to save this matrix to secondary storage would take days on commodity hardware. With increasing quality of data, the memory requirements for the pre-processing step in itself can grow exponentially. Apache Spark, a framework often used for processing very large data sets, can address the exponentially growing memory needs of this algorithm. We built a system using an MCL Spark library, that can bring down the processing time of the algorithm from days to a matter of minutes. We also built a scalable framework that can execute MCL and provide the necessary data to visualize the results. This system has a programmable web API, which users can leverage to run MCL and visualize the results on their own data sets.