

Search for single patterns is important and challenging. Signal search can be split into cases: Is the signal known? Is it partially known? Do we accept error? If we accept error, then what kind of error? Substitutions or also insertion-deletions of single letters? If we don't know it, we are searching for common instances of an unknown signal and the same questions pertain concerning errors. Searching for patterns can be complicated by the structure of the human genome since it makes it relevant to search conditionally on features in the human genome, such as "500 base pairs before a protein gene" or similar criteria. We could also have a battery of signals we are searching for simultaneously and it could be that the search could be rationalized relative to first searching for one signal, then the second etc. . To all these search scenarios are additionally added the question of statistical significance: is the observation (found configurations of signals) statistically significant? It could it as well have been observed in a random string.