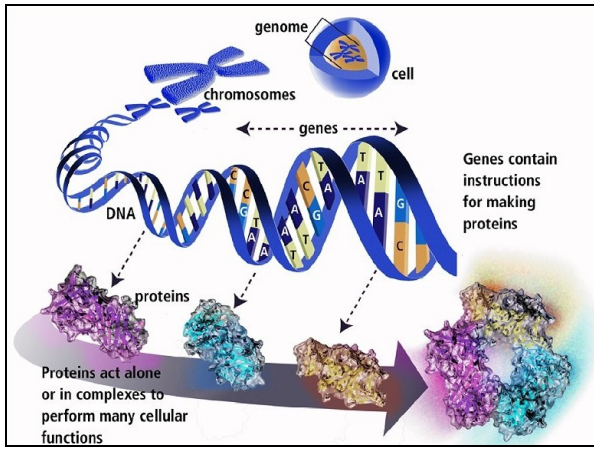
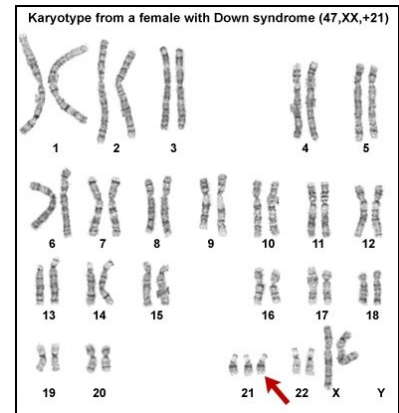


What engineers can do for biology and medicine?

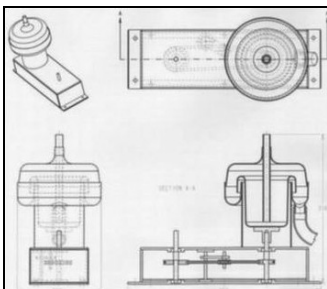
The adult human body contains billions of cells, each one has a different role. However, they all contain the same 23 pairs of chromosomes in their nucleus. The chromosomes are very long strings of DNA that harbor the genes: short DNA fragments that contain the code needed to create the proteins. In total, more than 20 000 genes are encoding for more than 40 000 proteins that are the real workers of the cells. Proteins are indeed responsible for many of the actions performed by the cells such as metabolism, cell division, and protein synthesis.



When one's genome is altered, that is when multiple copies of the same chromosome are present in the cells, there is a high risk of disease. A well known example is the Down syndrome, caused by the presence of three copies of the chromosome 21. Another example is the William syndrome caused by the presence of a single copy of the chromosome 7. The simple explanation is that more copies of a gene induce more copies of the corresponding protein which disturbs the very fragile system of the cell. Sometimes, the alteration is smaller than a whole chromosome and only encompass a few dozens of genes. While some genes, when altered, give rise to benign symptoms (or no symptom at all), some key genes are always causing very serious diseases when altered. Nowadays, we have defined more than 3000 diseases that are thought to be caused, at least partially, by chromosomal alterations. For some of them, we even know the causative genes, while some remain unexplained or partially unexplained. Thus, a major question in biology and medicine is: given a disease, can we find new genes that would cause that disease when altered ?



To answer that question, we have developed ENDEAVOUR. This software uses what is already known about the disease to predict which other genes might also be involved in that disorder. To make



accurate prediction, it uses different data sources that describes the function of the genes. These data sources can be seen as different views on the same object (i.e., gene) like an engineering drawing. By combining several incomplete data sources using various machine learning algorithms, ENDEAVOUR can make powerful predictions that can be further experimentally validated by biologists and geneticists.