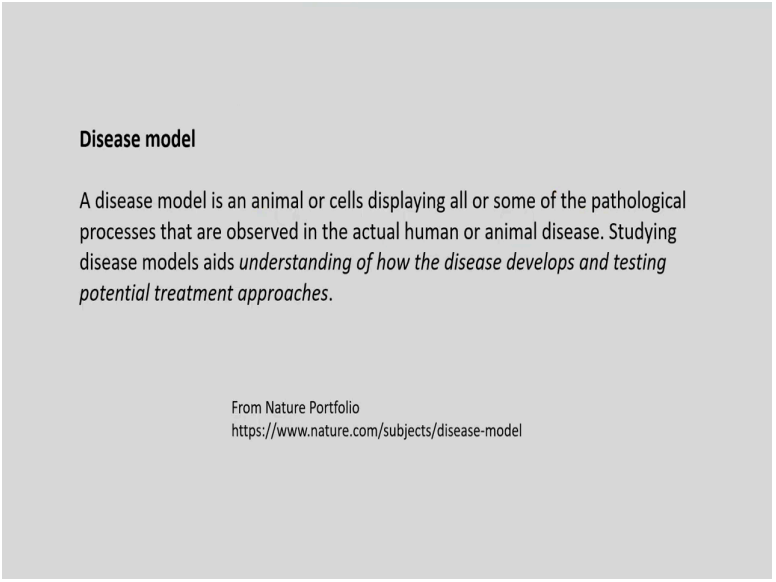
Genome Editing and Engineering
Prof. Utpal Bora
Department of Bioscience and Bioengineering
Indian Institute of Technology, Guwahati

Module - 09
Genome Engineered Disease Modelling
Lecture - 01
Animal models - Part A

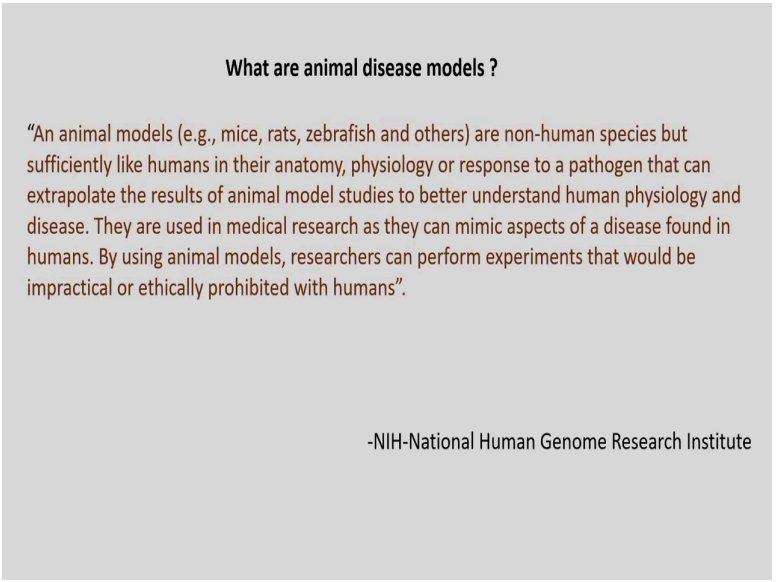
Welcome to my course on Genome Editing and Engineering. Today we are going to discuss about, module 9, where we will discuss about Genome engineered Disease models. In this module we will have discussions on cancer disease models, IPSC models and animal models. Let us start with, the Animal models which are quite old and well known.

(Refer Slide Time: 00:57)



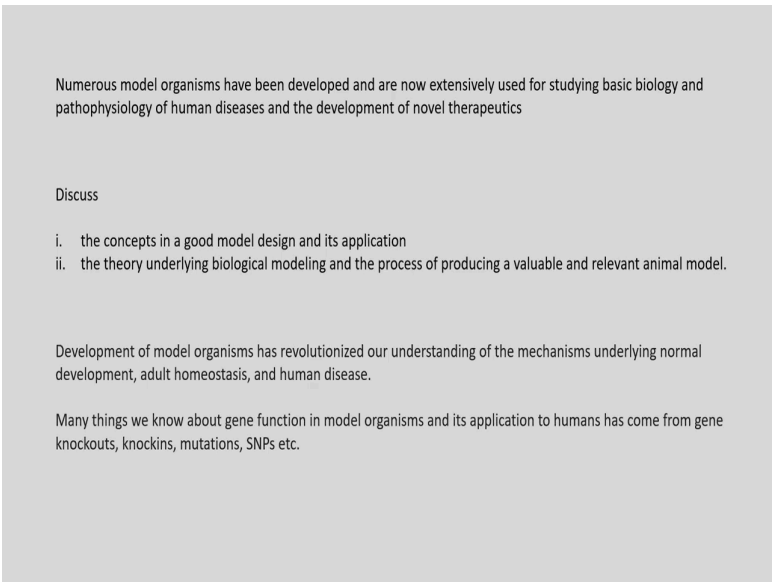
Before that, let us have a small discussion on disease model. What is a disease model? As per this definition from nature portfolio, a disease model is an animal or cells displaying all or some of the pathological processes that are observed in the actual human or animal disease. Studying disease models aids understanding of how the disease develops and testing potential treatment approaches.

(Refer Slide Time: 01:29)



So, National Human Genome Research Institute has defined animal disease models as below, an animal model which may be mice, rat, zebrafish and others are non-human species. But sufficiently, like humans in their anatomy, physiology or response to a pathogen that can extrapolate the result of animal model studies to better understand human physiology and disease. They used in medical research as they can mimic aspects of a disease found in humans. By using animal models, researchers can perform experiments that would be impractical or ethically prohibited with humans.

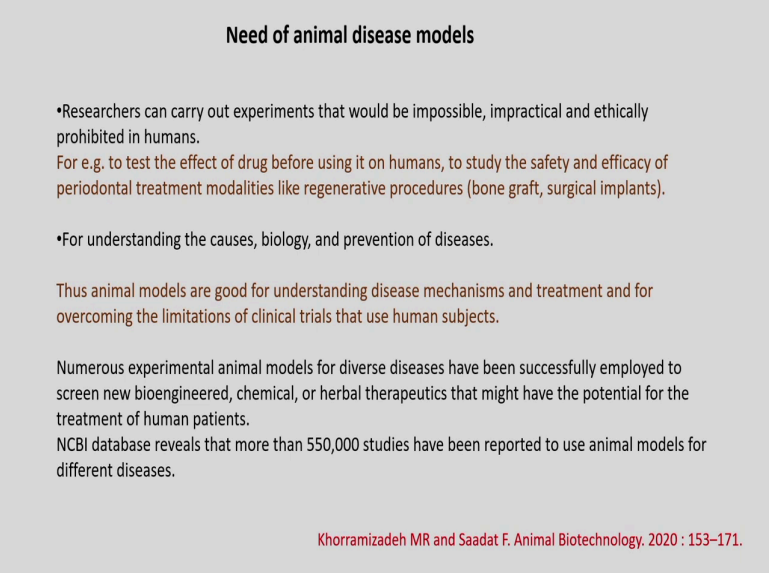
(Refer Slide Time: 02:17)



There are numerous model organisms which have been developed and are now extensively used for studying basic biology and pathophysiology of human disease and development of novel therapeutics. In our lecture, we will try to discuss and understand some of the concepts in a good model design and its application. And, the theory underlying biological modelling and the process of producing a valuable and relevant animal model.

Development of model organisms has revolutionized our understanding of the mechanisms underlying normal development, adult homeostasis and human disease. Many things today we know about gene function in model organisms and its applications to human has come from gene knockouts, knock-ins, mutations and artificially created SN Ps etcetera. And, in this regard the various genome editing technologies are very very helpful.

(Refer Slide Time: 03:21)



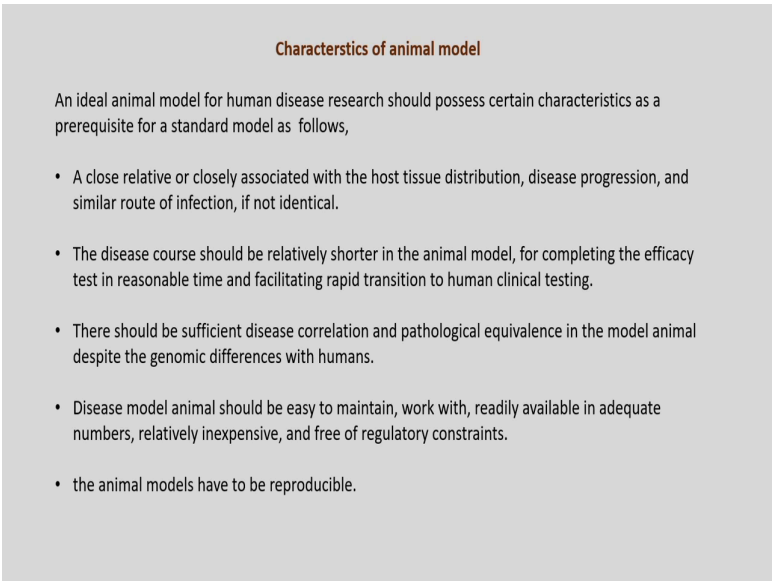
A question arises, why do we need animal disease models? As already told, it is impractical and ethically not permissible to carry out experimentation in humans. So, we have to go for alternatives like animals, but we have to create a disease model in the animal to further the experiments and also testing of drugs and therapeutics.

Researchers can carry out experiments as already discussed that would be impossible impractical and ethically prohibited in humans. For example, to test the effect of drug before using it on humans to study the safety and efficacy of some treatment modalities like regenerative procedures and for understanding the causes, biology and prevention of diseases.

These animal models are good for understanding disease mechanisms and treatment and for overcoming the limitations of clinical trials that use human subjects. Numerous experimental animal models for diverse diseases have been successfully employed to screen new bioengineered, chemical or herbal therapeutics that might have the potential for the treatment of human patients.

NCBI database reveals more than 550000 studies to be reported for use of animal models for different diseases.

(Refer Slide Time: 04:56)



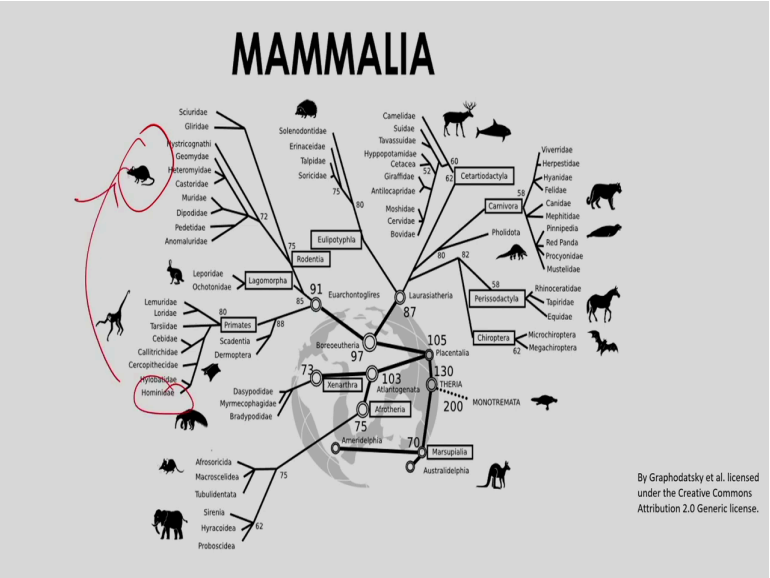
Now what are the characteristics of a good animal model? An ideal animal model for human disease research should possess certain characteristics as a prerequisite for a standard model as follows. A close relative or closely associated with the host tissue distribution, disease progression and similar route of infection if not identical.

Secondly, the disease course should be relatively shorter in the animal model for completing the efficacy test in reasonable time and facilitating rapid transition to human clinical testing. There should be sufficient disease correlation and pathological equivalence in the model animal despite the genomic differences with humans.

Disease model animals should be easy to maintain and work with, they should be readily available in adequate numbers for certain statistical designs. And this would be relatively

inexpensive and also free of regulatory constraints to the maximum extent possible. But there will be certain regulatory compliances the researcher has to follow and oblige.

(Refer Slide Time: 06:18)

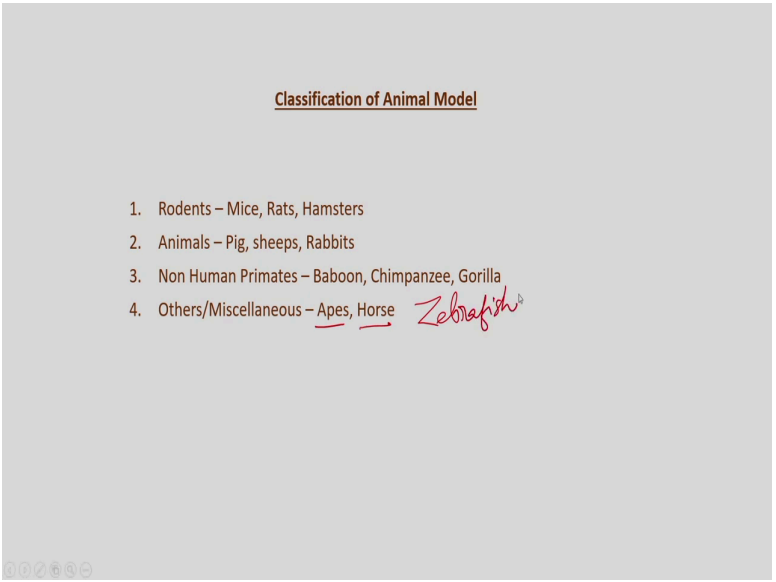


The most important thing is the animal models have to be reproducible. Now this is a kind of a relation you can see in the amongst the Mammalia and you can see the various branches here, then you have one important branch over here the primates under which you have Hominidae, under which you can find Homo sapiens. And, so these are their close relatives. Ideally, these closed relatives would make the most ideal animal model.

However, there are certain regulations which many a times prohibit the use of certain wild animals which may be endangered and other ethical reasons. Therefore, we have to go for animals which may be a little bit distantly related to the human species. But, there are other reasons as well we have to be able to have as already discussed in the earlier slides. These animals in large numbers for statistical analysis and the ease of handling is also one of the important things.

So, theoretically every animal can be actually converted into a human disease model. But, we usually do not go for using each and every animal we only select a few those animals we will be discussing in some of the future slides.

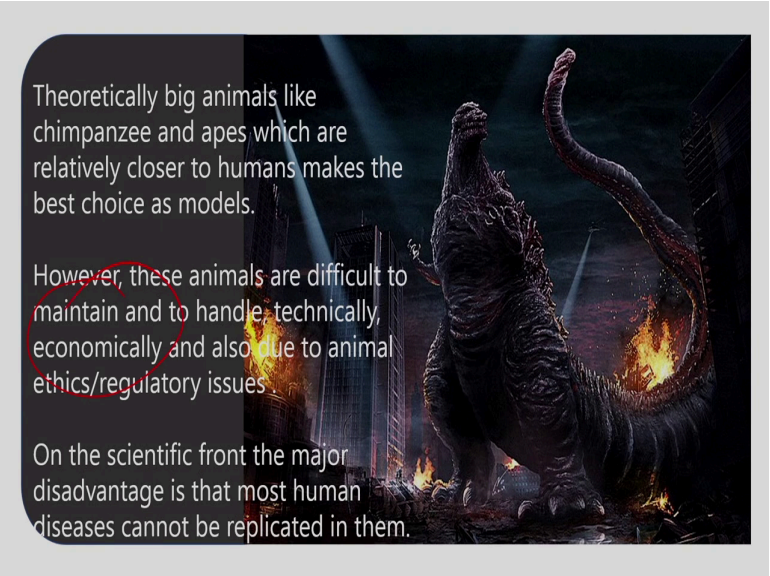
(Refer Slide Time: 08:02)



For example, we can classify the animal models into these main 4 groups; rodents, then animals, but it does not mean that rodent is not an animal. Here animal is a larger animal which is bigger in size than the rodents, under which the mice, rats and hamsters come. And, in larger models animal models we use the pig, sheep's, rabbits.

Then we have non-human primates like baboon, chimpanzee and gorilla and still very unrelated or miscellaneous like apes, horses and also in certain cases we use zebra fish. And, this list is not limited, there are many other animals you may find people using this is just to give you an idea about the classification of the animal models that we use for drug discovery or understanding certain molecular mechanisms.

(Refer Slide Time: 09:08)

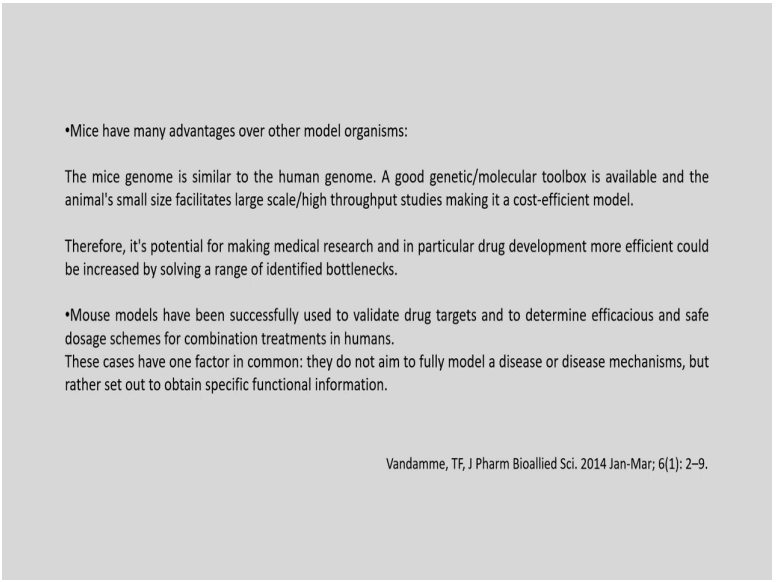


So, we have already discussed about the possibility of using large animals. But, these difficult animals are difficult to maintain and also economically very very expensive handling these big size animal. And, in from the scientific point of view one of the major disadvantage is that most human disease cannot be replicated in many of the animal models whether small and big.

(Refer Slide Time: 09:41)

The European Commission considered that, mice appears to be the most common genetically engineered animal model to study new drug development for different diseases.

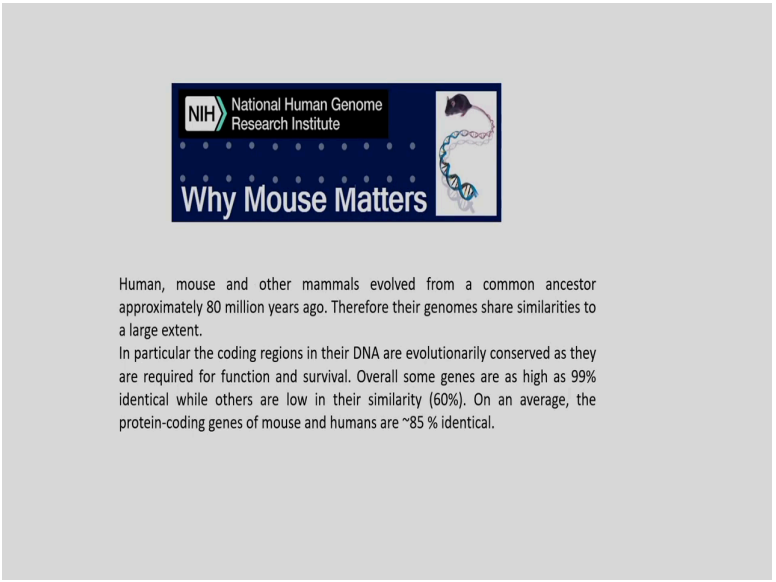
(Refer Slide Time: 09:55)



For numerous reasons such as, they are the mice genome is similar to the human genome. A good generic molecular toolbox is available and the animals small size facilitates large scale high throughput studies making it a cost efficient model. Therefore, its potential for making medical research and in particular drug development, more efficient could be increased by solving a range of identified bottlenecks.

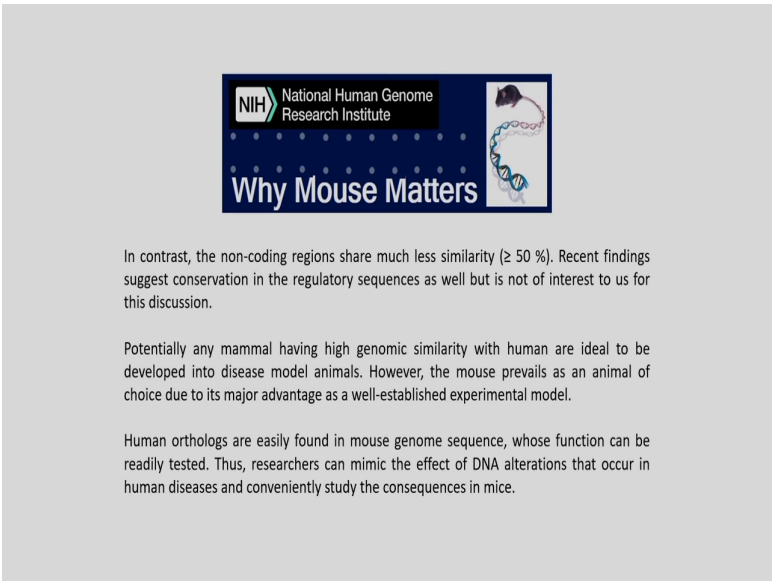
Mouse models have been successfully used to validate drug targets and to determine efficacious and safe dosage schemes for combination treatments in humans. These cases have one factor in common, they do not aim to fully model a disease or disease mechanism, but rather set out to obtain a specific functional information.

(Refer Slide Time: 10:47)



We know that human, mouse and other mammal's have evolved from a common ancestor; largely approximately around 80 million years ago. Therefore, the genomes are similarities to a large extent, in particular the coding regions in the DNA are evolutionarily conserved, as they are required for function in survival, overall some genes are as high as 99 percent identical while others are low in their similarity, on an average the protein coding genes of mouse and humans are around 85 percent identical.

(Refer Slide Time: 11:23)

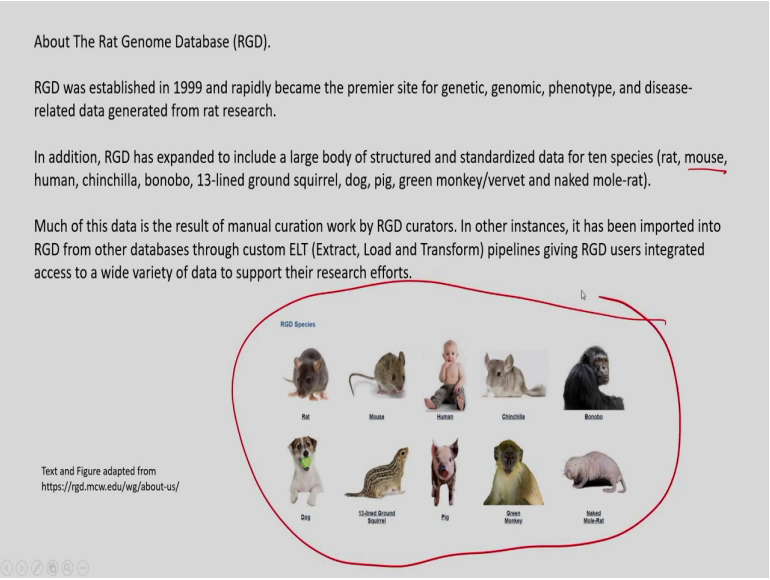


In contrast, the non-coding regions share much less similarity, less than 50 percent. And, recent findings suggest conservation in the regulatory sequences as well, but it is not of interest to us for these particular discussion.

Potentially any mammal having high genomic similarity with human are ideal to be developed into disease animal models; however, the mouse reveals as an animal of choice due to its measured advantages and well-established experimental method. Human orthologs are easily found in the mouse sequence, whose function can be readily tested.

Thus, researches can mimic the effect of DNA alterations that occur in human diseases and conveniently study the consequences in mice.

(Refer Slide Time: 12:12)



Another animal of choice is the rat, we have a rat genome database which you can assess in the website rgd dot mcw dot edu slash wg. And, these RGD was established in 1999 and soon it became a premier site for genetic, genomic, phenotype and disease related data generated from rat research.

RGD has expanded to include a large body of structured and standardized data for other species including mouse, a human, chinchilla, bonobo, squirrel, dog, pig monkey, vervet and mole rat and much of these data is the result of manual curation work by RGD curators. In other instances, it has been imported into RGD from other databases through custom extract load and transform pipelines.

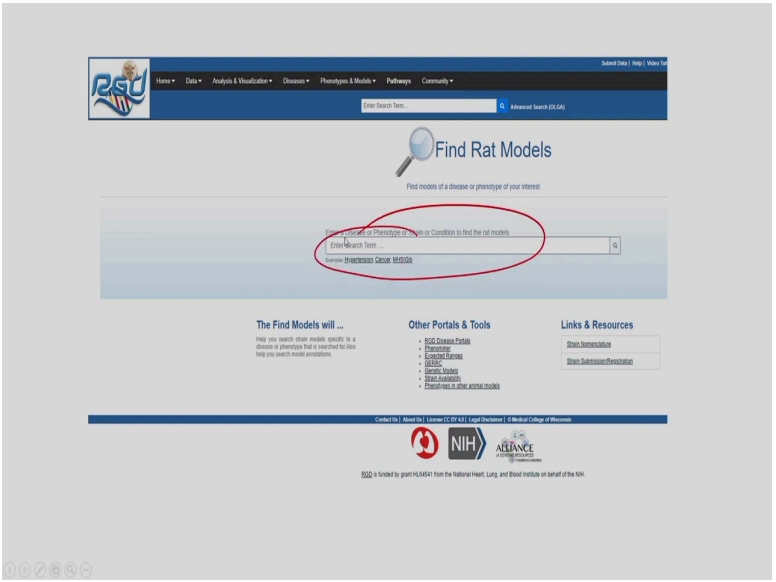
Giving RGD users integrated access to a wide variety of data to support their research. And you can see here the pictorial representation of the various RGD species or the species about which the genetic, genomic, phenotypic and disease related data are available in this database.

(Refer Slide Time: 13:48)

So, you can see here the web page of this RGD and if you go to the disease menu, you can see a lot of information regarding various diseases like age related cancer or cardiovascular, even COVID-19 data is now available including others like diabetes, liver disease, neurologic, respiratory and so on.

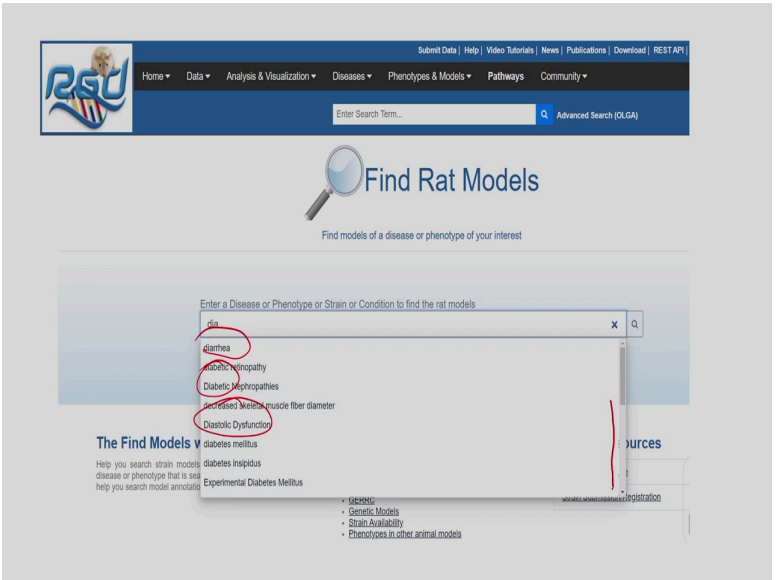
And you have many phenotypes and models also in this database and if you go into the drop down menu, you can see genetic models, autism models, then phenotypes and even you can find out models by typing the model you are looking for and this is a new feature in this website.

(Refer Slide Time: 14:47)



So, when you go to find new models, you can see here you can have to enter a disease or phenotype or strain condition to find the rat model over here.

(Refer Slide Time: 15:02)



And for example, I just type here the 3 u letters dia d i a dia. So, you, I have options in the drop-down menu forming things like diarrhea, diabetes, diastolic dysfunction and so on. And, these you can scroll down this drop down menu and you can find out the various disease conditions for which you have information regarding phenotype, strain or condition in rat models.

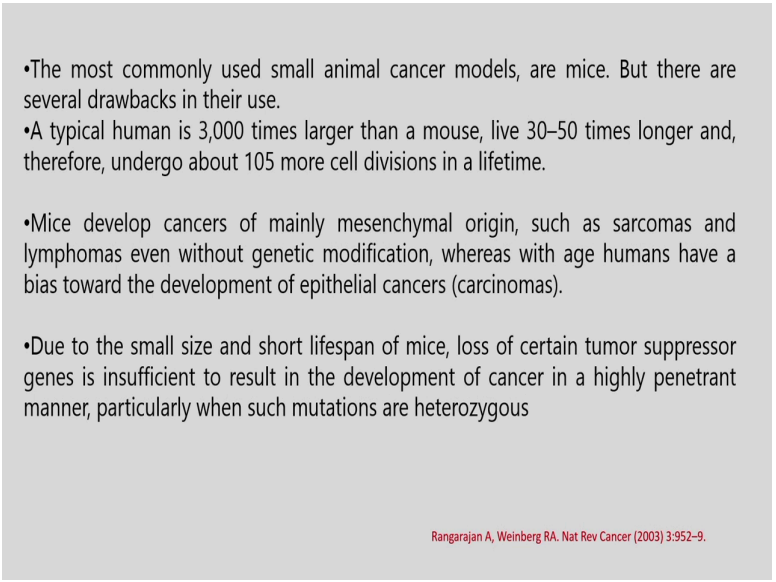
(Refer Slide Time: 15:31)

Other models which are becoming popular are, the pig disease model or porcine disease model and you can identify here a very famous personality Graig Venter who is now trying to develop humanized pigs, because it has lot of applications in diabetes, Parkinson's and even tissue engineering and disease of the heart, kidney, liver, small bubble and lungs.

So, here one of the aim here is to develop, these various organs and try to make them compatible with the human organs. So, in this regard a lot of genome editing work is being carried out to make those gene sets in a particular tissue, organ very very similar to human. So, that is the frontier area of research as regards animal models and animal disease models.

Apart from the ones for understanding diseases and finding out therapeutics, the tissue engineering and organ engineering is now becoming a very very important area of focus. Let us get back to the main discussion of animal models. The most commonly used small animal cancer models are mice. But, these are having several drawbacks such as a typical human is 3000 times larger than a mouse and leave 30 to 50 times longer and therefore, undergo about 105 more cell divisions in a lifetime.

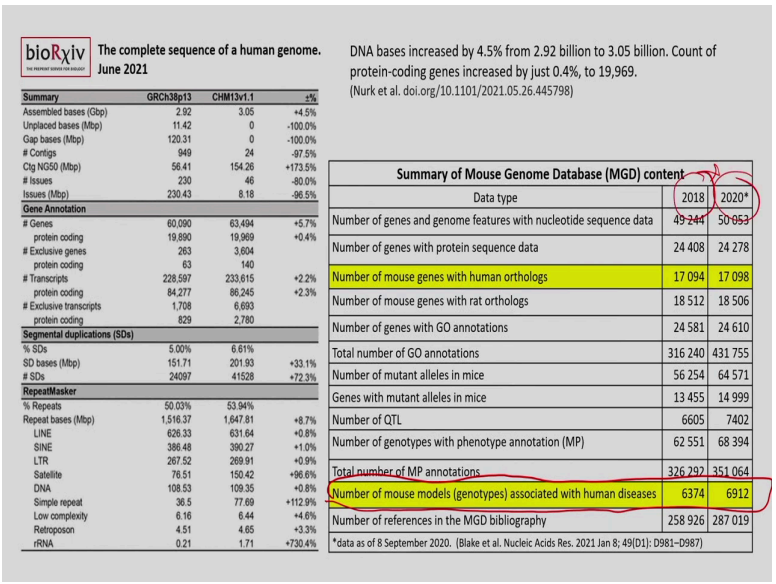
(Refer Slide Time: 17:28)



Mice developed cancers of mainly mesenchymal origin, such as sarcomas and lymphomas. Even without genetic modification whereas, with age humans have a bias towards the development of epithelial cancers. Due to the small size and short lifespan of mice, loss of certain tumor suppress genes is insufficient to result in the development of cancer in a highly penetrant manner particularly when such mutations are heterozygous.

So, when we are using mice models, we have to be careful in the data handling and data processing and their interpretation with respect to the human diseases.

(Refer Slide Time: 18:14)



Summary	GRCh38p13	CHM13v1.1	±%
Assembled bases (Gbp)	2.92	3.05	+4.5%
Unplaced bases (Mbp)	11.42	0	-100.0%
Gap bases (Mbp)	120.31	0	-100.0%
# Contigs	949	24	-97.5%
Ctg NG50 (Mbp)	56.41	154.26	+173.5%
# Issues	230	46	-80.0%
Issues (Mbp)	230.43	8.18	-96.5%
Gene Annotation			
# Genes	60,090	63,494	+5.7%
protein coding	19,890	19,969	+0.4%
# Exclusive genes	263	3,604	
protein coding	63	140	
# Transcripts	228,597	233,615	+2.2%
protein coding	84,277	86,245	+2.3%
# Exclusive transcripts	1,708	6,693	
protein coding	829	2,780	
Segmental duplications (SDs)			
% SDs	5.00%	6.61%	
SD bases (Mbp)	151.71	201.93	+33.1%
# SDs	24097	41528	+72.3%
RepeatMasker			
% Repeats	50.03%	53.94%	
Repeat bases (Mbp)	1,516.37	1,647.81	+8.7%
LINE	626.33	631.64	+0.8%
SINE	386.48	390.27	+1.0%
LTR	267.52	269.91	+0.9%
Satellite	76.51	150.42	+96.6%
DNA	108.53	109.35	+0.8%
Simple repeat	36.5	77.69	+112.9%
Low complexity	6.16	6.44	+4.6%
Retroposon	4.51	4.65	+3.3%
rRNA	0.21	1.71	+730.4%

Summary of Mouse Genome Database (MGD) content

Data type	2018	2020*
Number of genes and genome features with nucleotide sequence data	49 244	50 053
Number of genes with protein sequence data	24 408	24 278
Number of mouse genes with human orthologs	17 094	17 098
Number of mouse genes with rat orthologs	18 512	18 506
Number of genes with GO annotations	24 581	24 610
Total number of GO annotations	316 240	431 755
Number of mutant alleles in mice	56 254	64 571
Genes with mutant alleles in mice	13 455	14 999
Number of QTL	6605	7402
Number of genotypes with phenotype annotation (MP)	62 551	68 394
Total number of MP annotations	326 292	351 064
Number of mouse models (genotypes) associated with human diseases	6374	6912
Number of references in the MGD bibliography	258 926	287 019

*data as of 8 September 2020. (Blake et al. Nucleic Acids Res. 2021 Jan 8; 49(D1): D981–D987)

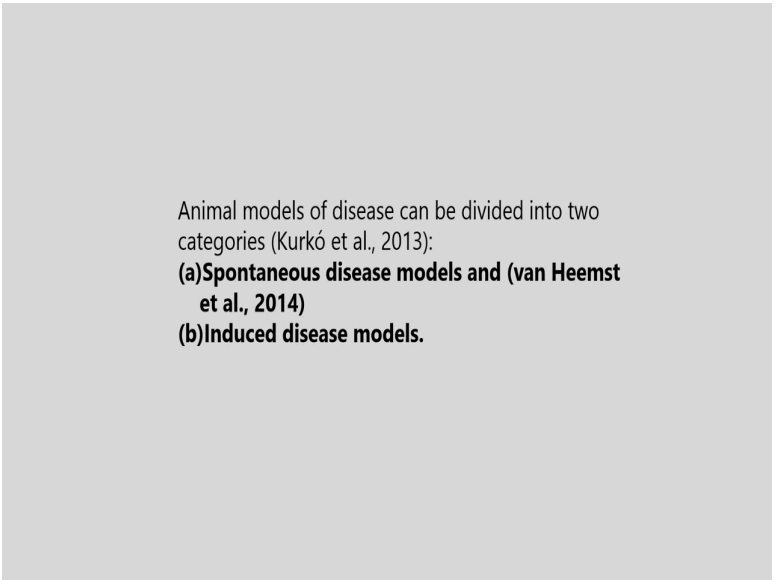
This is a comparison of the complete human sequence, human genome sequence completed in 2012.

So, prior to earlier reports the DNA bases have increased by 4.5 percent from 2.92 billion to 3.0 billion of because of the advances in and the **thrust** in the sequencing work. And, the count of protein coding genes of course, has not increased much it is hovering nearly around 20,000.

This is a summary of the similar statistics on mouse genome and a comparison between the human and the mouse. So, you have the total number of genes and genome features here it has increased from 49 to say 50,000. What is important for us to understand is the number of mouse genes with human orthologs something roughly around 17,098.

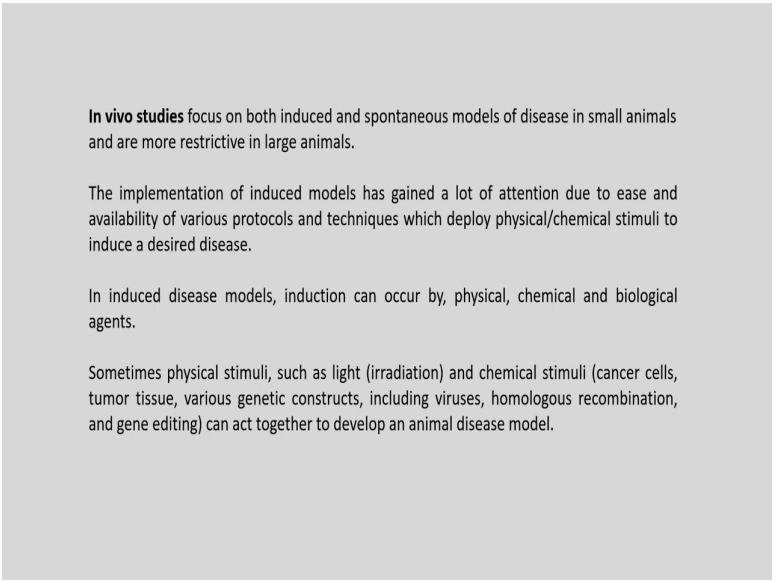
And, then, the number of mouse model genotypes associated with human disease has increased in the last 2 years from 6,374 to 6,900 by around roughly around 600 numbers. So, we can see the thrust of using mouse models for developing human diseases. So, this turns out to be roughly around 300 per year.

(Refer Slide Time: 20:12)



We can divide animal models of disease into 2 categories; the number 1, is the spontaneous disease models and the 2nd one is the induced disease models.

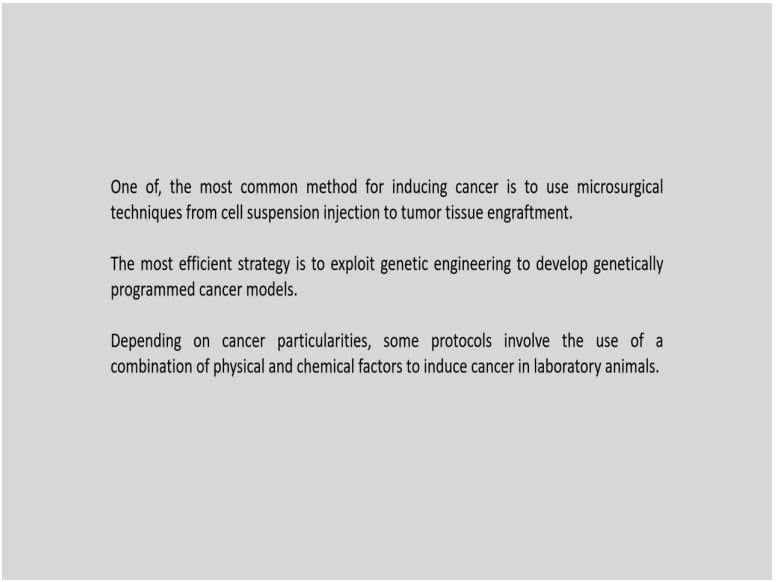
(Refer Slide Time: 20:23)



In-vivo studies focus on both induced and spontaneous models of disease in small animals and are more restrictive in large animals. The implementation of induced models has gained a lot of attention, due to ease and availability of various protocols and techniques which deploy physical, chemical stimulus to induce a desired disease.

In induced disease models, instruction can occur by physical, chemical and biological agents. Sometimes physical stimuli such as light and chemical stimuli connect together to develop an animal disease model.

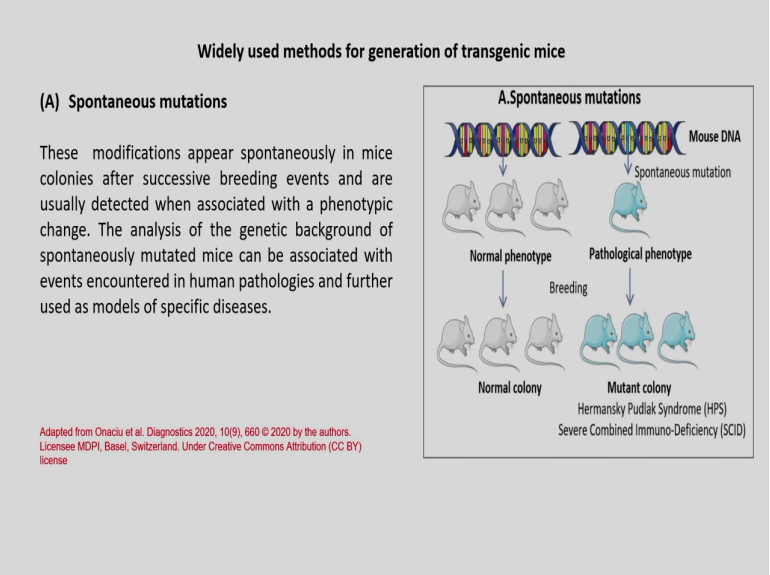
(Refer Slide Time: 21:05)



One of the most common method for inducing cancer is to use microsurgical techniques from cell suspension injection to tumor tissue engrafting.

The most efficient strategy is to exploit genetic engineering to develop genetically programmed cancer models. Depending on, the cancer particularities some protocols involve the use of a combination of physical and chemical factors to induce cancer in laboratory animals.

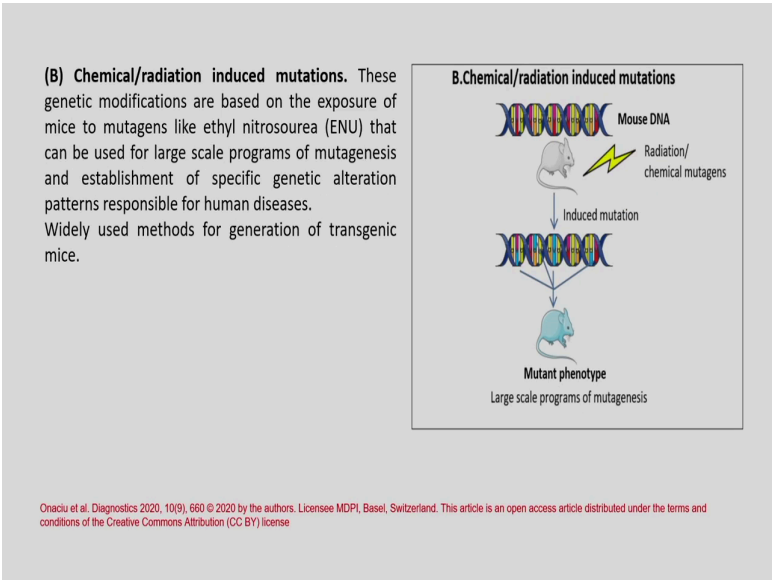
(Refer Slide Time: 21:33)



Let us discuss about some of the widely used methods for generation of transgenic mice. And many of these can be used for generation of disease models. Number A is the spontaneous mutation, you can see here the normal type and due to a spontaneous mutation, we get a pathological phenotype for example, and as a result of the breeding and in breeding we may get normal colony and we may we may get a mutant litter or mutant colony.

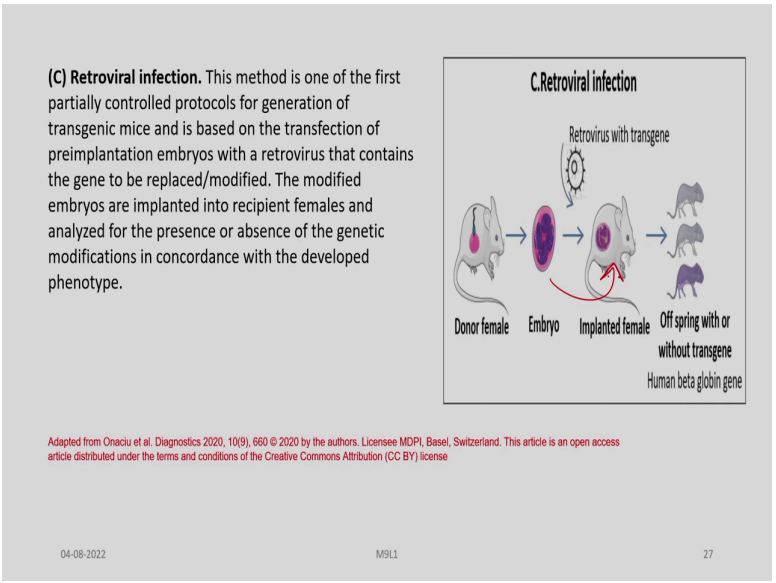
So, these spontaneous mutations as the name suggests, appear spontaneously in mice colonies after successive breeding events and are usually detected when associated with a phenotypic change. The analysis of the genetic background of spontaneously mutated mice, can be associated with events encountered in human pathologies and further use as models of specific diseases.

(Refer Slide Time: 22:45)



The 2nd case is about chemical or radiation induced mutations. So, we are using here either radiation or some chemical mutagens which induce mutations and results in a mutant phenotype. Here, genetic modifications are based on the exposure of the mice to mutagens like say ethyl nitrosourea and there are many more such agents. And this can be used for large scale program of mutagenesis and establishment of specific genetic alteration patterns responsible for human these days.

(Refer Slide Time: 23:27)

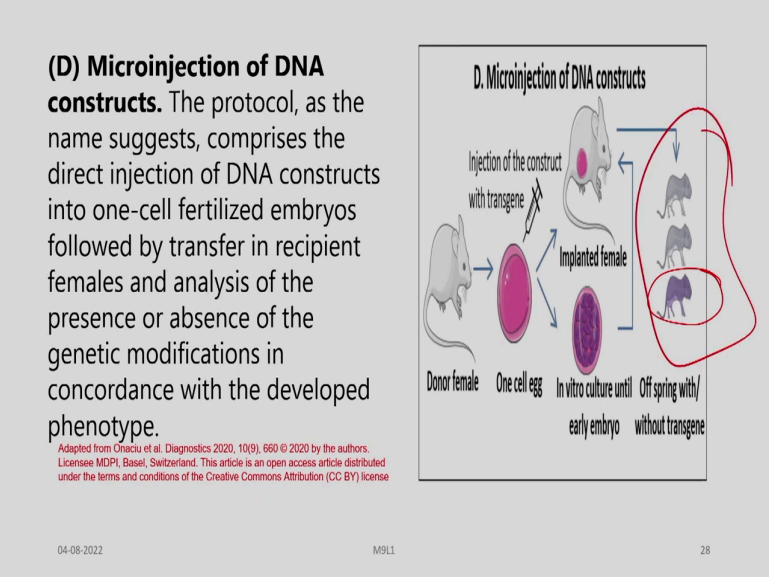


This is a widely used method for generation of transgenic mice. Number 3 is retroviral infection. This method is one of the first partially controlled protocols for generation of transgenic mice and is based on the transfection of pre-implantation embryos with a retrovirus that contains the gene to be replaced or modified.

The modified embryos are implanted into recipient females and analyzed for the presence or absence of the genetic modifications in concordance with the developed phenotype. And we can use these very easily for genome editing using the three known technologies that we have discussed ZFN, TALEN and CRISPR Cas-9.

So, here as you can see in this figure, there is a donor female from which the embryo is taken and it is infected with a retrovirus which is having a transgene and then it is re-implanted back into a mother and then we get offsprings with or without transgenes in this case they have experimented with the human beta globin gene.

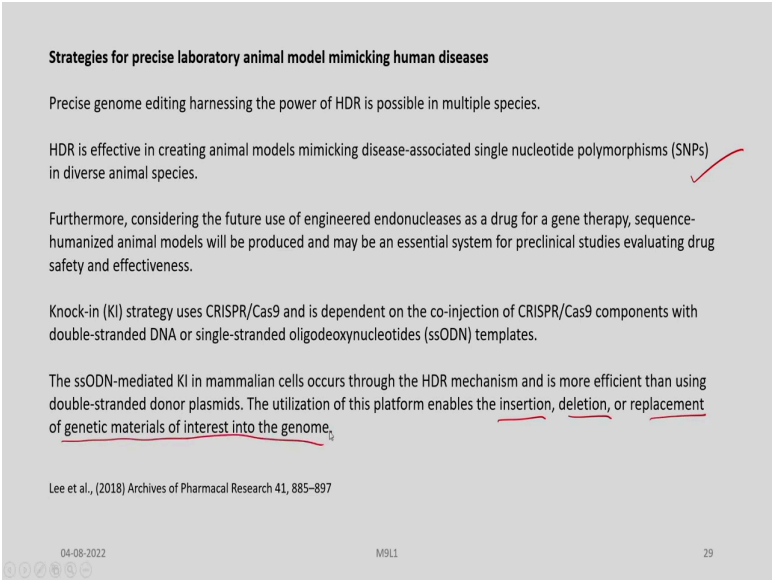
(Refer Slide Time: 24:44)



The 4th type of manipulation is through micro injection of DNA constructs and as you can see here a donor female and this is a one cell egg and the injection of the construct with transgene is done here and is implanted into the female which results in offspring with or without transgenes. So, this is with the transgene shown here schematically.

So, as already discussed, the system comprise of direct injection of DNA constructs into one cell fertilized embryos followed by transfer in recipient females and analysis of the presence or absence of the genetic modification as shown here in the picture.

(Refer Slide Time: 25:33)

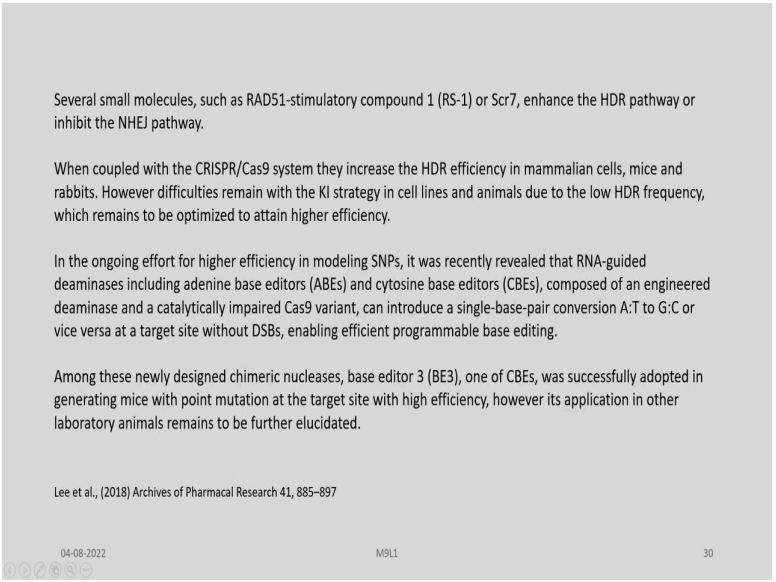


Let us now discuss in brief about strategies for precise laboratory animal models for mimicking human diseases. Precise genome editing harnesses the power of homology directed repair, which can be used in a multiple species. HDR is effective in creating animal models mimicking disease-associated single nucleotide polymorphisms in diverse animal species.

Considering the future use of engineered endonucleases as a drug for gene therapy, sequence humanized animal models will be produced and may be an essential system for preclinical studies evaluating drug safety and effectiveness in the future. In knock-in strategy CRISPR Cas-9 is used and it is dependent on the co-injection of CRISPR Cas-9 components with double stranded DNA or single stranded oligo-nucleotides as templates.

The single stranded oligo-nucleotide mediated knock-in in mammalian cells occur through the homolog directed repair mechanism and has been found to be more efficient than using the double stranded donor plasmids. The utilization of these platform enables the insertion, deletion or replacement of genetic materials of interest into the genome.

(Refer Slide Time: 27:08)

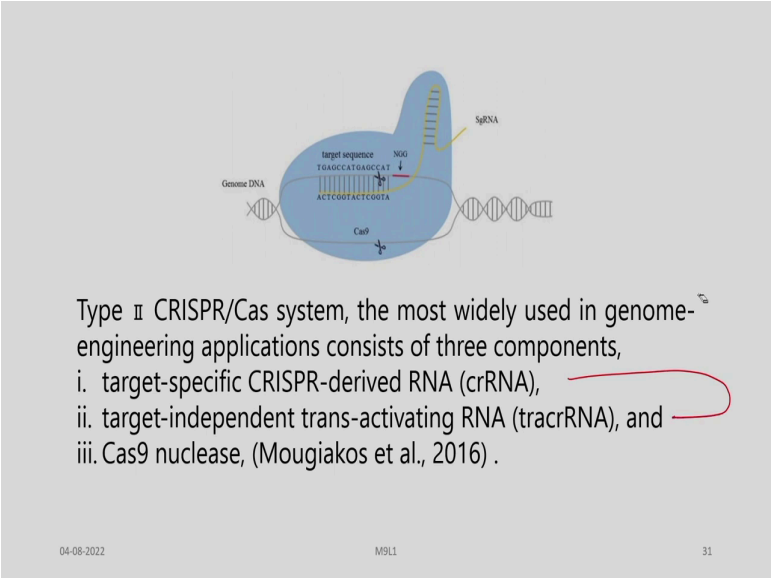


There are many small molecules for example, RAD51 stimulatory compound 1, RS-1 or SCR-1 which enhances the HDR pathway or inhibit the non-homology end joining pathway. When coupled with CRISPR Cas-9 system these compounds increase the HDR efficiency in mammalian cells, mice and rabbits; however, difficulties remain with the knock-in strategy in cell lines and animals due to the low homology directed repair frequency which remains to be optimized for attaining higher efficiency.

In the ongoing effort for higher efficiency in modeling SNPs it was recently revealed that RNA-guided deaminases including adenine-based editors and cytosine-based editors composed of an engineered deaminase and a catalytically impaired Cas-9 variant can introduce a single base conversion A to T or G to C or vice versa at a target site without double strand brakes enabling efficient programmable-based editing.

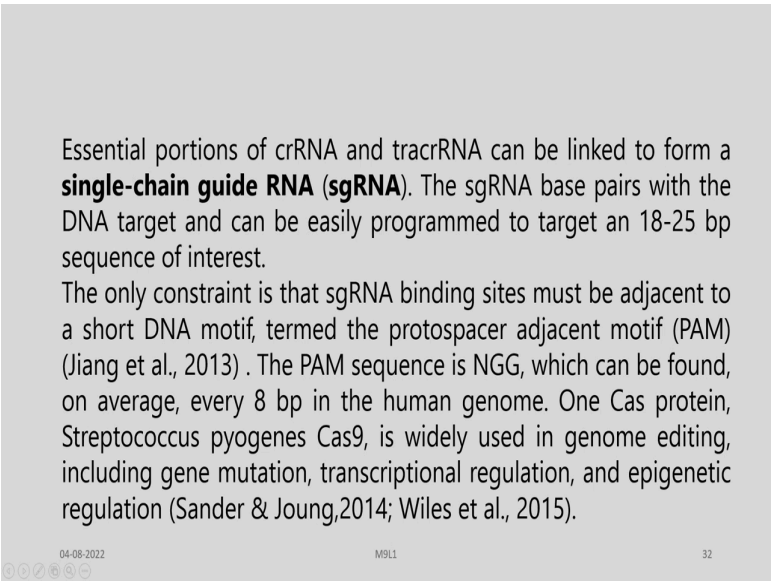
You have already learnt about the various dead Cas-9 which are being converted into base editors in earlier lectures. Among, these newly designed chimeric nucleases base editor 3 or BE3 was successfully adopted in generating mice with point mutation at the target site with high efficiency; however, its application in other laboratory animals are yet to be elucidated.

(Refer Slide Time: 28:53)



You are familiar with this CRISPR Cas system let us not spend time in discussing about the various components the CRISPR RNA, tracer RNA and the Cas-9 nucleus and the joining of these two components through a linker to generate the single guide RNA.

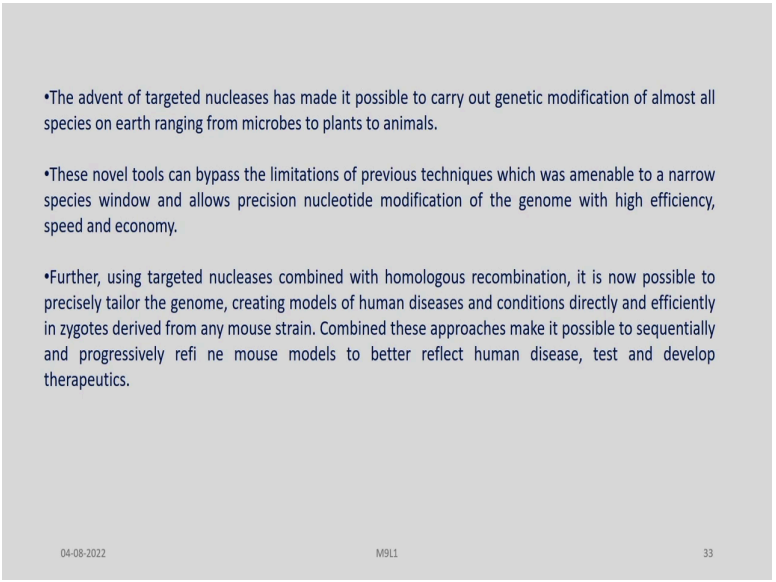
(Refer Slide Time: 29:16)



And we know the sg RNA base pairs with the DNA target and we can program it to target an 18 to 25 base pair sequence of interest and we know about the importance of the PAM sequences having the NGG signature and the role of Cas-9 which is the nuclease and also

apart from being a cutter the modified Cas-9 can be converted into other applications like transforming it into a best editor and so on.

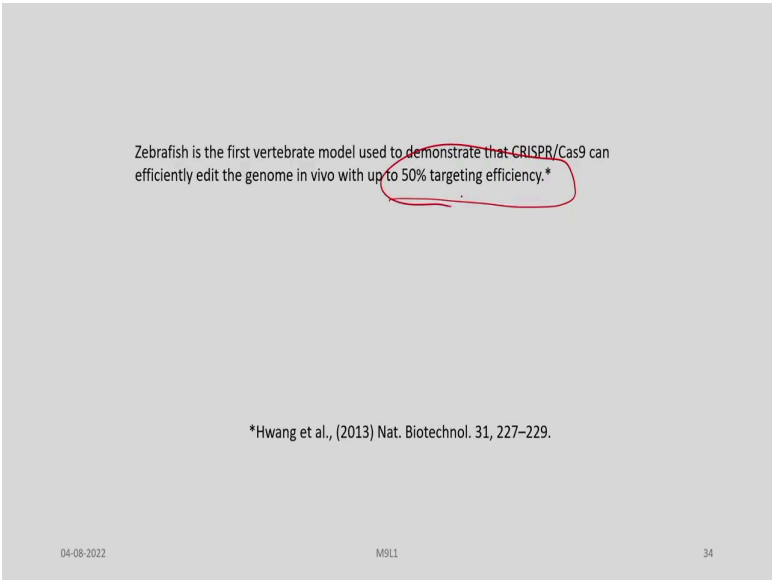
(Refer Slide Time: 29:59)



Overall, the advent of targeted nucleases has made it possible to carry out, genetic mutations of almost all species on earth ranging from microbes to plants to animals and these novel tools can bypass the limitations of previous techniques which was amenable to a narrow species window and allows precision nucleotide modification of the genome with high efficiency speed and economy.

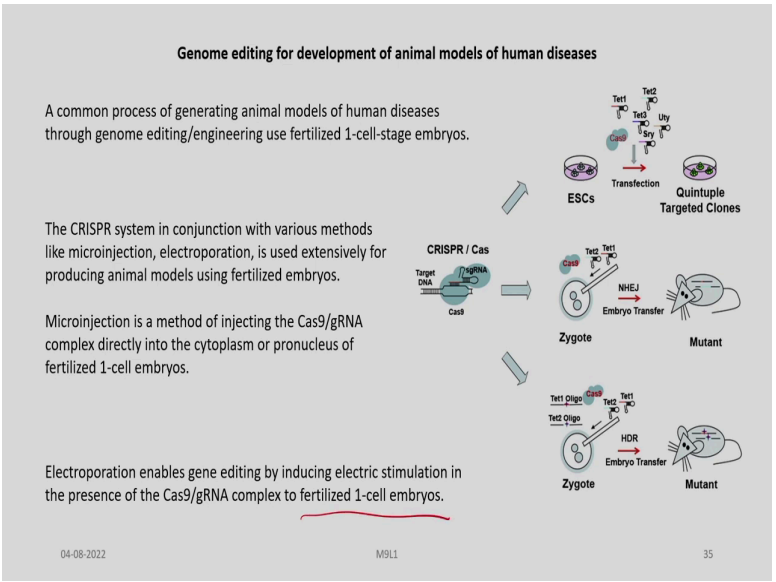
Using targeted nucleases, combined with homologous recombination, today it is possible to precisely tailor the genome creating models of human diseases and conditions directly and efficiently in zygotes derived from any mouse or animal strain. Combined approaches make it possible to sequentially and progressively refine these models to better reflect human disease test and develop therapeutics.

(Refer Slide Time: 30:56)



In this regard, we need to mention that zebrafish is the first vertebrate model used to demonstrate that CRISPR Cas-9 can efficiently edit the genome in-vivo with up to 50 percent target efficiency which is quite high.

(Refer Slide Time: 31:18)



So, this is a brief overview of the genome editing for development of animal models of human diseases.

A common process of generating animal models of human diseases through genome editing uses fertilized one cell stage embryos. The CRISPR system in conjunction with various methods like micro-injection, electroporation is used extensively for producing animal models using fertilized embryos. Micro-injection is a method of injecting the Cas-9 gRNA complex directly into the cytoplasm of pro-nucleus of fertilized one cell embryos.

While electroporation enables gene editing by inducing electric stimulation in the presence of Cas-9 gRNA complex to fertilized one cell embryos. With this we come to an end of Part-A of Animal Models. We will continue our lecture in Part-B of this module.

Thank you.