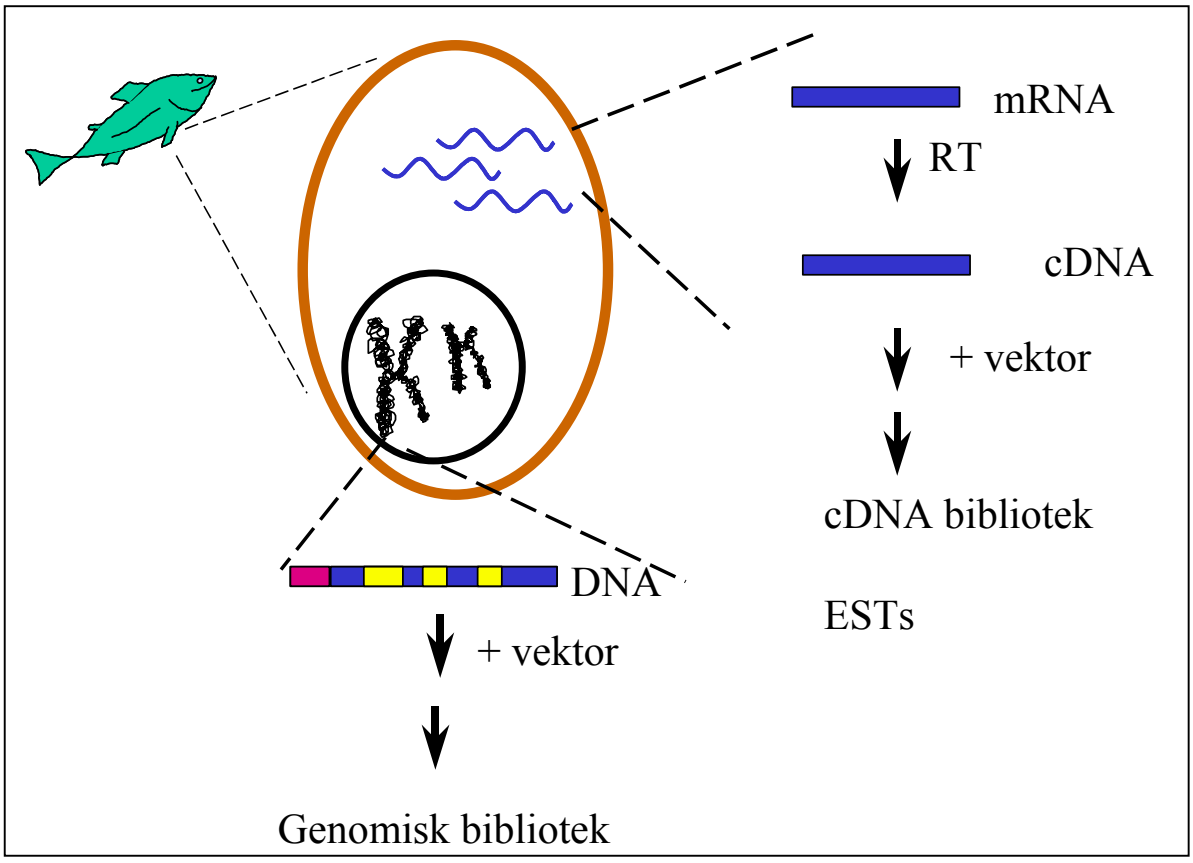
CodGen:

Genetic mapping and Functional biology in Atlantic cod.



Proposal for a research programme:

CodGen - genetic mapping and functional biology in the Atlantic cod (*Gadus morhua*)

A proposal for a national research programme on genetic mapping and functional biology in Atlantic cod is drawn up by the Institute of Marine Research (IMR), the University of Bergen (UiB), Norwegian Institute of Fisheries and Aquaculture Research (NIFAR), the Norwegian College of Fishery Science (NFH) at the University of Tromsø and the National Institute for Nutrition and Seafood Research (NIFES). The total budget is NOK 150 million over 5 years.

A working group appointed by the Ministry of Fisheries in a letter dated August 13, 2002 has been the main authors. Its members are: Ole J. Torrisen, IMR; Geir Lasse Taranger, IMR; Audun Nerland, IMR; Steinar Bergseth, Norwegian Research Council (NFR); Gro-Ingunn Hemre, NIFES; Trond Ø. Jørgensen, NFH; Atle Mortensen, NIFAR and Ivar Lossius, UiB.

Other members of staff of the participating institutions have also helped to draw up the proposal. We would like to mention in particular the following persons: Johan Lillehaug, Gunnar Nævdal, Sigurd Olav Stefansson, Arild Folkvord, Geir Dahle and Frank Nilsen.

Ole Torrissen

3 May 2003

1. Summary

The main objective of this project is:

- To establish structural and functional genomic knowledge in Atlantic cod in order to solve biological problems associated with management and harvest of wild cod stocks and development of a substantial aquaculture production of cod in Norway.
- Bring the competence in molecular biology up to a high international level for the fishery research institutions of Norway (Institute of Marine Research (IMR), the University of Bergen (UiB), Norwegian Institute of Fisheries and Aquaculture Research (NIFAR), the Norwegian College of Fishery Science (NFH) and the National Institute for Nutrition and Seafood Research (NIFES))

The Atlantic cod is one of our most important fish resources, and the species also has the potential to become a farmed species of significance in both national and international terms. The rapid technological advances being made in molecular biology and bio-informatics will allow the genome of the cod to be mapped within a few years, and they also open up the possibility of large-scale functional genome studies of cod. The genome of the cod appears to be simple and compact (approximately 920 million base pairs) in comparison with e.g. salmon. The mapping and functional genome studies on cod will benefit both the fishery and the aquaculture industry.

We propose a five-year national programme of research with a budget of at least NOK 150 million to map the cod genome and to carry out functional genome studies. Such a programme would make it possible to develop a number of products and technologies based on sequencing information from cod, as well as help us to understand and exploit cod as a cultivated species and an important fish resource. The fisheries research centres in Bergen and Tromsø are capable of carrying out a project of this sort and, just as importantly, of making use of the information that it generates. The programme will be based on the technological platforms, which have been established via the FUGE programme, as well as on strategic programmes at the participating institutions.

2. Justification of the proposed programme

The human genome is now spelled out to the last letter, polished and wrapped up, and a range of species are partially sequenced or in the line to be sequenced. The genomes of mouse, fruit fly, nematode worm, thale cress, yeast, rat, African clawed frog, pufferfish and zebra fish are nearly finished or already completed. New candidates in the line are the honeybee, chicken, chimpanzee, kangaroo, dog, pig and silkworm (Holmes, 2002. New Scientist 21/21 Dec. 2002).

The panel at US the National Genome Research Institute use the following points in prioritising the suggested genome projects:

1. The organism must be of practical importance, and an immediate application is needed.

2. The second main criterion is how hard the task will be. All else being equal, a short genome should be favoured over a long and complicated one.

2.1 Why cod?

The cod full fill the criteria put up by the panel at US the National Genome Research Institute completely, by both having an immediate need for the knowledge and being a relatively small genome. Furthermore, the 5 institutions behind the proposal have the facilities and human resources needed to coordinate and implement the project. The cost of the project is also reasonable in relation to the benefits for the marine environment, wild stocks and the emerging cod aquaculture industry.

The need for genomic knowledge of cod is urging within three principal areas:

- ***Management of wild cod stocks.***
- ***Monitoring impacts of anthropogenic environmental factors and climatic changes***
- ***Removing biological bottlenecks in cod aquaculture.***

Management of wild stocks:

- Effect of changes in population sizes and fishing pressure on genetic structure and phenotypic characteristics.
- Recruitment success
- Identification of stocks
- Tracing of cod in catches and in the market
- Establish the genetic baseline for Norwegian cod stocks.

Monitoring impacts of anthropogenic environmental factors and climatic changes:

- Effect of climatic changes on recruitment and performance of cod in Atlantic waters.
- Effect of disposals from oil industry on sexual maturation and recruitment of cod
- Effects of domestic pollutants on survival and performance of cod.
- Deformities and diseases in farmed and wild cod stocks.
- Interactions between farmed and wild cod

Removing biological bottlenecks in cod aquaculture:

- Understanding better the mechanisms behind unexpected high mortalities during start feeding and early juvenile production

- Controlling sexual maturation in farmed cod in order to increase productivity.
- Develop strategies for feeding cod to enhance product yield, growth, feed conversion and product quality.
- Understanding the mechanisms behind deformities in order to prevent the excessive deformities in cod farming.
- Establish sensitive indicators for environmental preferences for farmed cod.
- Achieve a better understanding of the unique immune system of cod, enabling development of effective vaccines and vaccine strategies against actual pathogens.
- Selective breeding for increased productivity by:
    - Develop rapid genotyping methods
    - To use Quantitative Trait Loci (QTL) studies to identify genomic regions for traits of interest
    - Develop extensive genetic maps to increase our possibilities of identifying QTL's and consequently give the basis for marker Assisted Selection.

The Norwegian Salmon Genome project started as a bioinformatic activity within different EU projects. In co-operation between the Norwegian Veterinary College and the Norwegian Research Council (NRC) this project was further developed into EST sequencing and bioinformatics project, and recently into a Functional Genomics Project.

The project group for CodGen is of the opinion that these two proposals should be handled independently, but NRC has insisted on a ranking between these two projects. The key question here is whether the industrial and governmental benefits would be higher by developing the Salmon Genome Project into a full genetic mapping of the salmon genome.

- The CodGen project group acknowledge the fact that 40 000 EST's (mainly from Canada) published in GenBank is an argument in favour for converting the Salmon Genome Project into a full genome mapping programme.
- However, the salmon genome is large and complex (basically tetraploid). A total mapping of the salmon genome would be extremely expensive. The kangaroo genome (approximately the same genome size as salmon) is estimated to cost between 700 and 1050 million NOK. The salmon genome would probably in total be even more expensive, probably up to 10 times the cost of the CodGen project.
- The driving forces in the salmon genome research are Canadian research laboratories. The Norwegian Salmon Genome Project is in danger of duplicating work already done in Canada.
- The salmon industry is a running industry served by multinational companies. The project would probably only have few short term benefits

exclusively for the Norwegian salmon industry or for the government. So far the industrial involvement in the project have been zero.

Table 1. Estimated genome size and costs of genome mapping programmes for some actual species.

Specie	Genome size (bill bases)	Cost [1] (million NOK)
Honeybee	0.270	49
Chicken	1.2	210
Chimpanzee	3.0	210-350
Kangaroo	3.3	700-1050
Dog	2.8	210-350
Silkworm	0.500	105
Cod	0.920	150
Atlantic salmon[2]	3.5-4.5[2]	1500

1) Cod and Atlantic salmon estimated by IMR, the others by New Scientist, 2002.

2) Considered essentially tetraploid, assembly would therefore be a tough task.

3. Background and justification of the proposed programme

The Atlantic cod (Gadus morhua), is widely distributed across the continental shelf regions of the North Atlantic. Several important cod stocks are of great economic and social importance. In Norwegian waters, the Norwegian-Arctic stock, which spends most of its life in the Barents Sea, is the most important. There are also a large number of more or less well-defined local stocks along the coast of Norway. Other important cod stocks are found in areas off Iceland, Greenland and Canada and in the North Sea and the Baltic. These have all suffered from extreme fishing pressure in recent years, and several of them are showing clear signs of overexploitation. Things have gone particularly hard with the Northern cod stock on the Grand Banks off the east coast of Canada. Until recently, this stock has been showing no signs of recovery even after ten years during which fishing has been banned, following the collapse of the stock in the early 90s. With poor catches as a result of reduced stocks in Norwegian waters as well, total catches of Atlantic cod have displayed a worrying downward trend during the past 10 - 15 years.

The cod is a well known and popular specie which has a large market world wide. The potential for cod farming is therefore regarded as extremely high on a global basis. Cod also enjoy a number of natural advantages as a potentially important cultivated species: a well known reproduction biology, short egg and larval stages, rapid growth, good feed utilisation, suitable behaviour (”tame”),

relatively immune to stress and, as far as we know, good health. Cod also appear to adapt well to traditional sea-cages, even to the extent that aquaculture technology developed for salmon can easily be adapted to cod. Most of the challenges offered by cod farming concern start feeding, cannibalism in the young fish phase, lipid deposition in the liver and premature sexual maturation. We can also expect to encounter challenges on the health side, particularly related to bacteria, viruses and parasites.

Cod may also be suitable for use as a model species for monitoring climatic change and the effects of environmental toxins on the marine environment. Differences in temperature-dependent recruitment have been demonstrated in different stocks. In the northern parts of its distribution range stocks have a tendency to display relatively better recruitment in relatively warm years, while the opposite is the case in southern stocks. How far this is due to simultaneously varying environmental conditions or is determined by genetic adaptations to variations in food availability is not known. However, there is little doubt that several cod stocks are to be found in water masses that are considerably colder than those are that provides maximum growth rates under good nutritional conditions. Recent studies carried out by the Institute of Marine Research indicate that alkyl phenols released by offshore platforms along with production water are capable of disturbing the hormonal balance and reproductive capacity of cod. This suggests that increased petroleum activity in regions in which cod grow may represent a threat to the species. Large quantities of pollutants are still being released into the sea, and we have only very limited knowledge of the effects on marine organisms of these tens of thousands of chemicals, whether they are created directly by man or result from human activities. The cod is a key species in sensitive northern regions, and is capable of being utilised as one of several model species in studies of the effects of various types of emissions from industry and the petroleum sector.

Our understanding of how life processes are controlled and of how organisms develop from embryo to adulthood has grown explosively in the course of the past decade. The cause of this increased insight has been highly dependant on the progress that has been made in molecular biological research and the use of computer science and technology to analyse large quantities of molecular biological data. While we formerly possessed detailed knowledge of one or a few genes from an organism, information is now available about the complete genomes of more than 50 bacteria species, yeast, the plant *Arabis (Arabidopsis) thaliana*, the zebra fish (almost complete), the fugu, drosophila, rice and the human being. The genome of several other species will soon be fully sequenced, so that several hundred thousand genes are known. In spite of such technical knowledge, relatively little is known about the function of individual genes or about possible functional interactions among genes. Functional genome research is still in its early infancy and will develop into the key to biological insight and tp understand the genetic relationships. Since proteins are the active genetic products, what is known as proteomics will play a key role in biological research in the future.

A series of comparative studies of individual cod stocks, including coastal cod, have demonstrated a significant degree of genetic variation in a number of important characteristics, such as rate of growth, energy distribution and age

at first sexual maturation. The stability of genetic markers has been discussed for thirty years. This is really a discussion of how rapidly evolution is capable of taking place. Some scientists claim that some allelic genes can undergo relatively changes in frequency in the course of a short period of time: even from one generation to the next. In non-functional markers (micro-satellites, etc.) the rate of mutation is so high that even these markers are not stable enough to characterise whole populations. A project that aims to map the genome, and at the same time is capable of detecting changes from one generation to the next, could at least shed light on the frequency of mutations, and possibly also on natural selection. The cod is the species about which we possess the best background information concerning genetic variation at individual loci, as well as being one of the few species on whose gene frequencies we have historical data. We assume that natural cod stocks are adapted to natural environmental conditions, and that this adaptation has taken place via a process of natural selection over a large number of generations. We still do not know the selection criteria involved, but the overriding aim of each individual of a species is to survive sufficiently long to pass on its genes to the next generation. Experience has shown that selection has not led to reductions in individual diversity among individuals, as there is a great deal of individual variation even among apparently well-adapted populations. In aquaculture there are no natural selection pressure and the fish are selected for traits of economical interest under farming condition. An efficient selective breeding programme for cod will primarily been based on a combined family- and individual selection, although progeny testing will also be considered. With the knowledge we have today it is not possible to do a selection based on genotype in cod since we do not know which genes or genomic regions are responsible for the different traits. If we, through studies of the genome, can identify regions at the genome responsible for important traits under farming conditions, selective breeding could be developed more efficiently.

Selective breeding based on QTLs implies that such loci (QTL) are linked to the trait of interest for selection. Extensive genetic maps will increase our possibilities of identifying QTLs and consequently give the basis for Marker Assisted Selection.

Functional genome studies of cod will represent an important scientific and practical approach to mapping the genetic basis of efficient aquaculture of this species, will help us to understand the processes that take place in natural populations, and enable us to demonstrate the effects of pollution and other human activities. At their most basic, biological processes are regulated by the expression of genes, which are subsequently processed and integrated into an individual response to an internal or external signal to the organism. The ontogeny of cod consists of phases of intense cell division and differentiation into various organs and types of tissue. Throughout the embryonic and larval phases there is a continuous process of development and differentiation of the musculature, skeleton, sensory system, digestive system, respiratory organs and circulatory system, not to mention the immune system. Sexual differentiation and the development of gonads also take place early in life. The cod does not develop its adult characteristics such as its musculature, sensory and nervous systems, digestive system and respiratory

system until metamorphosis. The new possibilities offered by large-scale genome research will allow the study of thousands of genes that are important for these developmental biology processes.

A precondition for efficient functional genome research is that large parts of the genome should already be known. Mapping of the genome of the Atlantic salmon has already begun at national and international level, but little has been done on cod in this area. Recent data suggest that the salmon genome is about four times as large as that of the cod, and the cod genome appears to have a very simple organisation in comparison with the salmonids. This suggests that it would be significantly simpler and quicker to carry out a cod genome mapping project than to complete the equivalent for salmon. Experience which has already been gained in the course of genome programmes on fish such as salmon, zebra-fish and fugu, as well as the bacterium *Methylococcus capsulatus*, will considerably simplify a cod project, and the economic return in comparison with the costs involved would be much greater.

We propose a research programme aimed at mapping the genome of the cod and using this knowledge in functional genome studies of processes that are of importance for cod farming, for understanding processes in wild populations, and for using the cod as a model species for understanding the effects of human activities on the marine ecosystem. We envisage the programme as being a national effort, in which the Institute of Marine Research (IMR), the University of Bergen (UiB), the Research Council of Norway, Norwegian Institute of Fisheries and Aquaculture Research (NIFAR), the Norwegian College of Fishery Science (NFH) and the National Institute of Nutrition and Seafood Research (NIFES) will play important roles. Other institutions should also be invited to contribute. The promoters of this proposal can point to a considerable amount of scientific competence and research facilities that would enable them to carry out such a programme of research. Their expertise and facilities include knowledge of the biology of the cod, both in natural conditions and in aquaculture, research stations and research vessels, a high level of molecular biological knowledge and infrastructure, as well as the bio-informatics competence that is essential for dealing with the huge amounts of data that will be generated. The University of Bergen has just completed a project on mapping the genome of the bacterium *Methylococcus capsulatus*, which was done by putting about 90% of the sequencing process out to tender. The University of Bergen and the Institute of Marine Research are members of the national salmon genome project, and are already pursuing a great deal of activity on the molecular biology of both salmon and cod. Moreover, the Institute of Marine Biotechnology at the Norwegian College of Fishery Science have studied various immune mechanisms and genes in the cod for years and the University of Bergen and the SARS Centre are also carrying out research on the zebra-fish as a model specie. The zebra-fish has been established as the most important model fish species, and is of particularly great importance in studies of developmental biology. Since the genome of the zebra-fish is about to bee fully mapped and genetic and molecular biology research are well established in Bergen, it will be a simple matter to establish synergetic contacts between a new cod project

and the research which is already being done on zebra-fish. Such links have already produced good results in salmon.

The Norwegian Institute of Fisheries and Aquaculture Research is now in the process of starting a national selective breeding programme on cod. They are also using gene technology methods (microsatellites) when studying different populations and studying relationship between individuals. The work on establishing DNA markers will be developed parallel with the quantitative selective breeding to make this work as efficient as possible.

Experience of genome mapping in other species (e.g. human beings) has shown that a conscious attitude to patenting and commercial exploitation of the data generated is essential. Applications of sequencing information are being patented in more and more connections, and Norway could risk that companies in other countries will seek patents on the use of gene sequences for nationally important species such as salmon and cod. This could involve genes that are important in combating disease, in breeding, in developing particular farming methods or identifying specific populations. Processes initiated by American companies are already under way to patent genes in salmon and cod that could be of great importance in aquaculture. It is also likely that a cod genome project would reveal a very large number of hitherto unknown genes, and there would be considerable elements of bio-prospecting in this process.

A cod genome project should be established as a collaborative venture involving several Norwegian research institutions and be carried out within a reasonable period of time. International participation in such a research network would also be advantageous.

The research data to be obtained by the project must benefit the society as a whole. Rights to the use the outcome of the results must be guaranteed via agreements between the project partners. This must be done in such a way that the public interests are served, while encouraging commercial/industrial development based on the results of the research. If the Norwegian partners wish to ensure that they can maintain optimum control of the commercial exploitation of the results of a cod genome programme, the project should be fully financed from Norwegian sources. External services (DNA sequencing, etc.) can be purchased as required for the successful completion of the project. This will also enable the project partners to make the results of the project public available if they decide to do so.

4. Proposed design and organisation of the research programme

Sub-programme 1: Sequencing and development of methods and tools for functional genome studies in cod.

4.1 Genomics

Objective: To obtain the best possible basis for carrying out modern basic and industry-related research in the areas of biology, ecology, genetics, molecular

biology and bio-informatics, where these are of relevance to the sustainable bio-production of cod.

a: Gene sequencing

According to the latest available information, the genome of the cod contains about 920Mbp (9×10^8 base-pairs).

The strategy for sequencing the genome will be a combined approach between whole genome shotgun sequencing and sequencing of BAC-clones. We will make two types of shotgun libraries by physical shearing into simple type of vectors such as pUC19. One shotgun library can be managed in average insert size of about 2kb, and another shotgun library around 10kb. For the second shotgun library, we will not sequence all clones, rather they will serve as bridging the contigs from the smaller insert clones. We will use BACs with the considerations of gap filling, repetitive sequences, and ease of retrieving important gene/genes or other functional units

A large portion of the project will be bioinformatics, where our Chinese partner has extensive experience. First of all, they will set up a dedicated database and corresponding LIMS to manage the libraries, clones, sequences, and related data. After sequencing, we would use an assembly tool, either phrap or TIGR assembler, depending on the need for the quality or length, to assemble the genome. They will annotate the genome with various methods and programs, including Genscan, FgeneSH, blast against human, mouse, fugu genes etc. Finally, they have solid experience in SNP discovery at genome scale. All these extracted information will be stored in the database, and also be delivered in flat files with formats defined by our group.

On the basis of offers from our partners in China ($1 per 500 base-pairs) the costs will range from $2.6 million to $10 million to sequence once, three times or five times respectively, through the complete genome. A minor cost for establishing a BAC library will come in addition to this. A single genome sequence is adequate as means of obtaining good information about gene structure, and to allow identification of genetic markers. Triple sequencing of the genome is sufficient to construct a decent genetic map for use in gene identification, partial annotation and to permit bio-informatics. Studies such as identification of new genes are to be carried out, to analyse exon-intron transitions and to provide information about the regulation of genetic expression and chromosome organisation. Five to eight sequencing repetitions are required for a statistically complete coverage of the genome. Single sequencing of the complete genome, in combination with EST sequences (expressed genes), is regarded as the method of preference to search for genetic markers and other genetic information of relevance for our understanding of the response of the cod to spawning and growth conditions, for example. Changes in stress reactions could also be correlated with genetic markers. Such studies will improve the basis for population monitoring and the possible definition of early intervention standards. All in all, sequencing will provide us with a very large collection of genetic clones, which will be of great use in molecular biology experiments such as micro-matrices for gene expression and DNA arrays for genetic analyses. Micro-matrices will be an

important tool in the study of genetic expression in various organs and tissues, stages of development and growth conditions.

b: EST sequencing

In order to establish gene sequence data capable of being related directly to the physiological functions of specific organs or developmental stages, a major effort in cDNA (EST) sequencing will be required. mRNA will be isolated from different tissues (pituitary, ovary, testis, liver, spleen, kidney etc) at different developmental stages and after different type of stimulations. The mRNA will be converted into cDNA and libraries established in phage/plasmid vectors. Clones will be picked randomly or screened for specific inserts and sequenced. For cloning and sequencing of genes with low expression, numbers, normalized libraries will be used in the same way. Presumably, about 5-10% of the total genome should be determined by this means. This will cost a total of somewhere between NOK 10 and 20 million.

c: Micro-matrices

In collaboration with NTNU, the Radium Hospital in Oslo and the University of Oslo, the University of Bergen has established a national network with the aim of making micro-matrix technology available to Norwegian research groups. Micro-matrices allow studies of the expression of several thousand genes to be carried out simultaneously on a chip. The network is now operational and it has applied for FUGE funding for a further development of joint service functions.

4.2 Proteomics

An integral part of the cod genome project proposed by us is the implementation of structural and functional studies of proteins in various organs and at different stages of development. In addition to enzyme studies we will utilise 2D electrophoresis combined with mass spectrometry. These techniques will make it possible to efficiently identify the proteins and genes that are active in the development and growth of the cod. In this connection, it will be a natural step to couple the project up to the proteomics and structural/function platforms which are currently being set up in Bergen and Tromsø as part of the FUGE programme.

4.3 Bio-informatics

The programme will require a fully developed bio-informatics platform:

1) Storage and administration of genome sequences and data based on sequences.
2) Access to software for assembling, gene searching, naming and similar analyses, in addition to expertise in the use of these technologies.
3) Via the University of Bergen's GABI project, this expertise has been established, as has the ability to organise and perform major projects of this sort. The following elements will be required:

- a database infrastructure dedicated to sequencing projects, with particular emphasis on eukaryotic organisms.

- a wider range of software for assembly, gene searching, annotation, phylogenetics and other related analyses.
- the development of an interface between the individual programmes, so that certain analyses can be carried out automatically.
- the development of new software, or the upgrading of existing packages in order to optimise these for the current project.

It will be a natural step to couple the proposed programme up to the bio-informatics platform which is currently being set up in Bergen as part of the FUGE programme.

Sub-programme 2: Functional genome studies in cod

This sub-programme will utilise techniques such as DNA fragment analysis, micro-matrix hybridisation, in-situ hybridisation, quantification of gene expression by means of rtQ-PCR, Northern Blotting, Rnase protection assay (RPA), large-scale protein mapping (proteomics) and functional protein studies combined with experimental systems during all life stages of cod. We will also utilise experimental material from different populations and families from the selective breeding programme established at NFRI, Tromsø. The sub-projects will be basic in nature, but they will attempt to identify applications in the direction of understanding variations in natural populations, selective breeding, improving aquaculture methods, including optimised vaccination etc., developing environmental monitoring methods (biomarkers) in addition to identifying potential bases for applications in biotechnology.

Potential sub-projects

1) Population genetics, structure and traceability of populations - development of genetic markers for simple stock identification and determining the degree of genetic exchange between stocks.
2) The functional significance of genetic differences between stocks - studies of the genetic basis of different life histories in individuals from different stocks.
3) Basic developmental biological studies and early development, and strain differences. Development of individual organs and systems:
    i. Sexual differentiation
    ii. Development of the muscular system
    iii. Development of the skeleton
    iv. Sensory system
    v. Digestive system
    vi. Respiration
    vii. Immune system
4) Metamorphosis - studies of its genetic regulation

5) Growth, regulation of growth, nutrition and quality - identification of genetic factors that directly or indirectly regulate metabolism
6) Puberty, reproduction and gametes, detailed studies of selected genes involved in those processes, both in different organs as well as in time.
7) The immune defence system, health and vaccine development
8) The genetic basis of behaviour - links between patterns of behaviour and the occurrence of specific genotypes and protein function
9) Selective breeding of important traits for farming QTL studies make it possible for genotypes with desirable growth and production characteristics
10) Climatic and environmental effects on wild populations - identification of adaptation of genotypes to different environmental conditions (counter-gradient variation)
11) Effects of environmental toxins (e.g. alkyl phenols in produced water from oil fields).
12) Studies of nutritional ecology and toxicology.
13) Effects of oxidative stress – development of genetic biomarkers for oxidative stress with qTR-PCR.

5. Budgetary and funding proposal

A budget of at least NOK 150 million over five years is foreseen.

Sub-programme 1

a. Genome sequencing, including BAC library: at least MNOK 40
b. EST and micro-matrices: MNOK 30
c. Bio-informatics: MNOK 10

Sub-programme 2

a. Functional genome research: MNOK 70

Funding proposal

1) Ministry of Fisheries: 120 MNOK between 2004 and 2008
2) Contribution from the participating institutions: 30 MNOK
3) Alternatively: contributions from the Norwegian Research Council (NRC); The Research Fund

Table 2. Financial plan:

	2003	2004	2005	2005	2007	2008	Total
Own contributions	5	5	5	5	5	5	30
Ministry of Fisheries/NRC		30	25	25	20	20	120

6. Organisation, including list of partners

A cod genome project must be established as a collaborative venture involving several Norwegian research institutions and to be carried out within a reasonable period of time. International participation in such a research network will also be advantageous. Most of the sequencing (about 90%) has to be put out to tender.

Project partners:

The Institute of Marine Research (IMR)

The University of Bergen (including the SARS Centre) (UiB)

National Institute of Nutrition and Seafood Research (NIFES)

Norwegian Institute of Fisheries and Aquaculture Research (NIFAR)

The University of Tromsø, Norwegian College of Fishery Science (NFH/UiTø)

The programme would be headed by a programme board which will have one representative from each participating institution. The Ministry of Fisheries should appoint the chairman of the programme board.

An international scientific advisory group should also be appointed, with a mandate to provide advices regarding priority tasks and to evaluate the scientific progress of the programme on an ongoing basis.

7. Publishing, patents and rights, commercialisation and spin-offs

The research results from the project must benefit the society as a whole. Rights to the use of the research data must be guaranteed via agreements between the project partners. This must be done in such a way that the interests of society as a whole are served while encouraging commercial/industrial development based on the results of the research. If the Norwegian partners wish to ensure that they can maintain optimum control of

the commercial exploitation of the results of a cod genome programme, the project should be fully financed from Norwegian sources. External services (DNA sequencing, etc.) will be purchased as required for the successful completion of the project. This will also enable the project partners to make the results of the project public available if they decide to do so.