Kingfish genetics: commercialisation strategies

Authors: Wayne Knibb, Abigail Elizur, Josephine Nocillado and Adam Miller
Project participants: Nguyen Nguyen, Erin Bubner, Rob Lamont, Jane Quinn, Daniel Powell, Pablo Sanchis

Project No. 2013/700

AUSTRALIAN SEAFOOD COOPERATIVE RESEARCH CENTRE

Final Report, April 2015

CONFIDENTIAL

Title: Kingfish genetics: commercialisation strategies

PRINCIPAL INVESTIGATORS: Prof. Abigail Elizur and A/Prof Wayne Knibb

ADDRESS: Sunshine Coast University, Locked Bag 4, MAROOCHYDORE DC, QLD 4558, AUSTRALIA

This project was conducted by Cleanseas Tuna Ltd. and the University of the Sunshine Coast

This project was funded by the Australian Seafood Cooperative Research Centre, Box 26, Mark Oliphant Building, Science Park, Adelaide, Laffer Drive, BEDFORD PARK SA 5042 and the Fisheries Research and Development Corporation (FRDC), Locked Bag 222, Deakin West, ACT 2600.

ISBN 978-0-9751913-4-7

The Australian Seafood CRC is established and supported under the Australian Government's Cooperative Research Centres Program. Other investors in the CRC are the Fisheries Research and Development Corporation, Seafood CRC company members, and supporting participants.

Office Mark Oliphant Building, Laffer Drive, Bedford Park SA 5042
Postal Box 26, Mark Oliphant Building, Laffer Drive, Bedford Park SA 5042
Tollfree 1300 732 213 Phone 08 8201 7650 Facsimile 08 8201 7659
Website www.seafoodcrc.com ABN 51 126 074 048

Australian Government
Fisheries Research and Development Corporation

An Australian Government Initiative
CRC AUSTRALIA

Non-Technical Summary

- Previously, we have implemented a selective breeding program for Yellowtail Kingfish (YTK) and today all production is based on genetically selected stock, which is up to 20% superior than the wild stock.
- This current project focuses on whether realized selection response was achieved and adds:
    - further families to the pedigree
    - facilities (broodstock tanks) to the program
    - data management and long term planning
    - new DNA microsatellite markers for pedigree assignment
    - measures the genetic improvement at up to 20%
- Today, CST is the only kingfish production company with a genetic program that has progressed beyond breeding the first generation yet maintained diversity. Present CST is in the unique position of having well advanced F_2s and soon will have F_3s in production.
- We have produced recombinant FSH and demonstrated it acts to promote gonadal development in YTK

Project Objectives:
In the original application:

Objectives

- To maximise the commercial benefit of the genetic program through development of a YTK fingerling export sector

- To calculate the commercial value of the genetic program with respect to wild stock in Australia and overseas and undertake a cost benefit analysis of the value of improved YTK sales nationally and internationally

- To identify a business case for viability of the genetics program based on benefit to CST in terms of production gains, hatchery efficiency and interstate and international fingerlings sales

- To fully Integrate the genetic program with industry operations and develop relevant IP protection methods

- To ensure genetic best practice: maximising selection response while minimising inbreeding

- To ensure long term commercial value and sustainability of the CST YTK genetic program by limiting inbreeding to 1% and incorporating suitable management procedures

OUTCOMES ACHIEVED OR PLANNED

- An enhanced genetic improvement program for YTK including inbreeding management at or below 1 % per generation and a reduced generation time (<3 years) enabling increased rate of genetic gain relative to the current program. This was achieved in that new inbreeding per generation is about 1%, males at entering into the breeding program earlier (3-4 years).
- An industry managed commercialized breeding program for YTK. This was achieved.
- A supply of genetically improved YTK nationally and internationally. There have been some international sales and would have involved some of the genetically improved stocks.
- Approaches for genetic protection of improved stocks. This involves restriction of the genetic diversity (number of families) sold at any one time.
- Approaches for the advancement of pubertal development and oral delivery of hormones

LIST OF OUTPUTS PRODUCED

- All fingerlings from CST are now based on the breeding program set up through the CRC.
- Trials of the selected fish show they outperform offspring from wild parents by up to nearly 20%.
- Peers reviewed papers on the genetics of YTK.
- Planning document including selection objectives, criteria, CST genetics inventory, future plans.
- Web site templates to a. store data from individual broodstock tanks, b. a link to pedigree records for each fish, c. a link to the overall pedigree in pedigree viewer.
- New broodstock tanks were added to the genetic program and new wild parents introduced into the pedigree.
- The number of high quality DNA microsatellite markers from a kingfish transcriptome almost doubled to 18.
- Males as young as 3 years used in the mating tanks, thereby reducing generation time and increasing the rate of genetic improvement.
- Value of the genetic program assessed.
- Technology for recombinant gonadotropins developed and applied.
- Technology for oral drug delivery trailed in model fish.

Abstract:

Use of genetic stocks coincides with profitability for CST: We have measured and shown that the rate of improvement is as high as nearly 20% per generation, and the roll out of the genetically improved fingerlings across the industry coincided with the transition of CST from *loss to profit*, a transition no doubt underpinned by use of the new genetic stocks. Soon CST will start deploying the F_3 fingerlings, where gains are expected to be between 10—20%, offering a further significant opportunity not only to sustain profitability but

extend on it. Arguably, in view of the nearly 20% gains measured, it is unlikely CST would have become profitable, or as profitable, without the genetically improved stocks. Also in view of the substantial magnitude of the genetic gains, and presumably, differences among families, future experiments on a range of topics at CST apart from genetics may need to control for the genetic effects both in terms of comparing historical data, and for contemporaneous trials where only a few families are being used because results can reflect genetic sampling of few special families rather than treatment effects.

Procedures, skills and capacities in place for long term sustainable genetic program: Over the years of this and former projects, CST hatchery staff have been trained in genetics and the genetic management of kingfish, and similarly USC learnt from CST to maximize genetic resources for maximal financial outcome for the company. Today, USC has built over time a unique set of corporate assets in genetics –skills, specialist expertise, procedures, knowhow, experience and wisdom that permit this new project to build, over new generations, on past successful work maximizing genetic gain for growth and conformity while minimizing inbreeding within the constraints of current commercial operations at CST. During the life of this new project new infrastructure and assets were added to the project, including new broodstock tanks, new families, new high quality DNA microsatellite markers to manage the pedigree and new equipment to assist in high-throughput analysis of DNA markers. Today, no other genetic program for kingfish comes within a small fraction of the achievements of the CST program: no other program has produced F_1 on a commercial scale that can be selected and used as broodstock, and where attempted, such as in WA, the pedigree was reduced to essentially one male and one female. At CST, ample genetic diversity has been achieved not only at the F_1, but also the F_2 and will also be at the imminent F_3 generation.

Accelerating genetic gain: During the life of this project, increasingly younger fish, especially males, were used successfully as broodstock, thereby reducing the net generation time and increasing the yearly rate of genetic gain.

Success: In a very short time of just several years, CST and USC have, sometimes by hard fought trial and error, worked out how to operate successfully over many generations a commercial breeding program for mass spawning kingfish. Today, almost miraculously in view of the risks and hurdles to solve, all of CST production is based on genetically selected F_2 fish, and this will move to the F_3 shortly. We should note that these achievements are well ahead of comparable efforts to build genetic programs in the Seafood CRC for abalone and barramundi, and in terms of industry adoption rates, even ahead of the oyster genetic breeding programs.

Today, CST has a genetic selection program completely without peer for kingfish in the world and one of the better organized and advanced aquaculture breeding programs in Australia. No one else has genetic parameter estimates for kingfish stocks, or breeding values established under commercial conditions, nor bred from selected F_1 fish to produce F_2 fish now grown out under commercial conditions, or shown they can continue to minimize inbreeding *OVER* multiple generations. Definitely no one else has shown they have

the wherewithal, skills, experience and planning to put it all together, line up all the ducks, and form a commercial breeding program. With diligence, organization and "stamina" over years from scientists and industry, there is no reason why CST cannot continue to enjoy this commercial advantage it has obtained, but build on it, generation after generation, to support not only their local growout, but also the market for their exported YTK fingerlings nationally and internationally.

ACKNOWLEDGEMENTS

This project was supported by the Australian Seafood Cooperative Research Centre, the Fisheries Research and Development Corporation, Cleanseas Tuna and the University of the Sunshine Coast.

Contents

1 Background

This project is a continuation of work from the Broodstock and genetic management of YTK Project No. 2010/768. In brief, the first project (2010/768) initiated a commercial breeding program of kingfish at CST, while the second (2013/700) built on the former to enhance aspects of the project like rate of genetic gain, infrastructure and levels of genetic diversity as well as hormonal manipulation trials for shortening generation times. Application to continue and extend the 2010/768 project was made through the CRC legacy program, where the CRC attempted to identify areas of previous work that were successful and had prospect, with some further support, to offer some enduring legacy from prior CRC investments. The application was successful and the project commenced in 2014.

2 Data management

This project has operated three levels or approaches to data management, progressing on a scale from ease of use (at farm), flexibility and simplicity to more advanced Access data bases requiring specialist staff/ consultants to manage and modify. It is our experience that no single package suits all the needs of a dynamic and evolving pioneering genetic program such as that for CST kingfish (or really any other evolving aquaculture genetics program). Indeed the idea that one data management system is sufficient and "one size fits all" is really overly simplistic, and, importantly, inflexible. Also there are no single data management systems yet built that can handle the complexity of mass spawning groups of fish, and they cannot deal with the matters of mate selection, inbreeding management and so on.

In a dynamic environment, hatchery staff need the flexibility of simple Microsoft excel work sheets. Scientific partners and hatchery operators need a web based data storage and retrieval system to lodge and access data, to visualize current operations and fish in broodstock tanks, to access and visualize the pedigree and breeding values, and to be able to review data extending over years and generations. Last, as operations stabilized, and systems become more routine, there is an opportunity to develop more sophisticated data entry forms, linkage of data sets and analyses through, for example, Microsoft Access data bases. It is likely there will be a requirement for all three levels of data management for the next few years.

2.1 Excel worksheets

This is the simplest form of data acquisition and storage. An example (TABLE 1) is the information collected on fish selected for size and condition. Here there is a requirement for:

- Each animal to be written on a single row
- Each animal to have a unique identifier
- Attributes and traits of the animal are written in consecutive columns

TABLE 1. Example of simple pedigree data

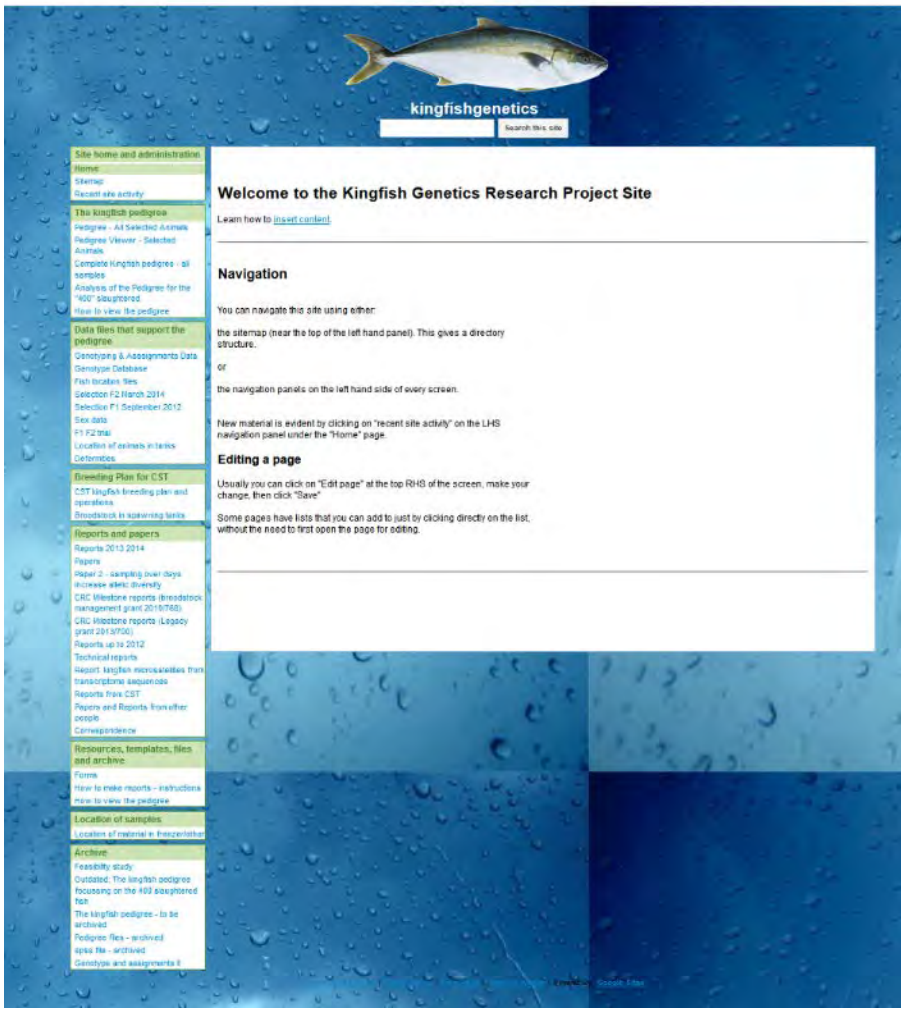
2011 selected broodfish 11/9/12				
Fish ID	Tube ID	PIT Tag	Weight	Length
2011-1	B1	168003	2.85	58.5
2011-2	B2	171160	3.1	57.5

2.2 Web based data storage, retrieval and analysis

2.2.1 General aspects

A dedicated kingfish genetic web site presents a rich environment to lodge and retrieve data, track versions and interact with team members via commenting (FIGURE 1). Currently our dedicated USC-CST site hosts about 90 pages of content relevant to the CST genetic program in easy to follow headings and menus. Many team members, at USC or CST, can edit their relevant pages using simple initiative procedures not available in most web sites. Different related pages can be linked. These sites have been in constant use for many years, so have withstood the test of time regarding adoption, ease of use, flexibility, and functionality.

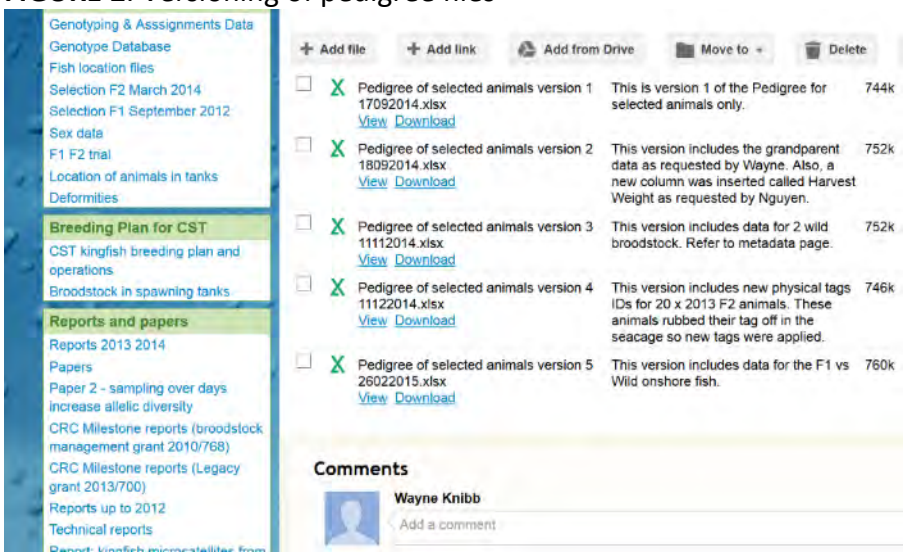
FIGURE 1. Home page of the kingfish website



2.2.2 Simple pedigree files

At its simplest, the web site host pages to the simple lodgement of data files, such as the pedigree at CST. The files are updated, but old versions left (FIGURE 2). This simple arrangement is superior to emails in various ways, first, there is little ambiguity over which file is the latest version, staff can comment of the file, and re-post edited versions etc.

FIGURE 2. Versioning of pedigree files



2.2.3 Data lodgement and retrival over many years.

A particularly powerful aspect of the web site is hosting files and correspondence that spans many years (as often is the case for genetic programs). Considering a specific example, a web page from the kingfish site hosts all details of a recent trial comparing the progeny of wild fish vs selected fish. Details of the initial planning, through the evolution of the plans, to experimentation, data analyses and finally report/ paper writing are all available, including over 100 emails, data records and reports, in chronological order and in various folders.

2.2.4 Visualization of breeding plan and linking pages

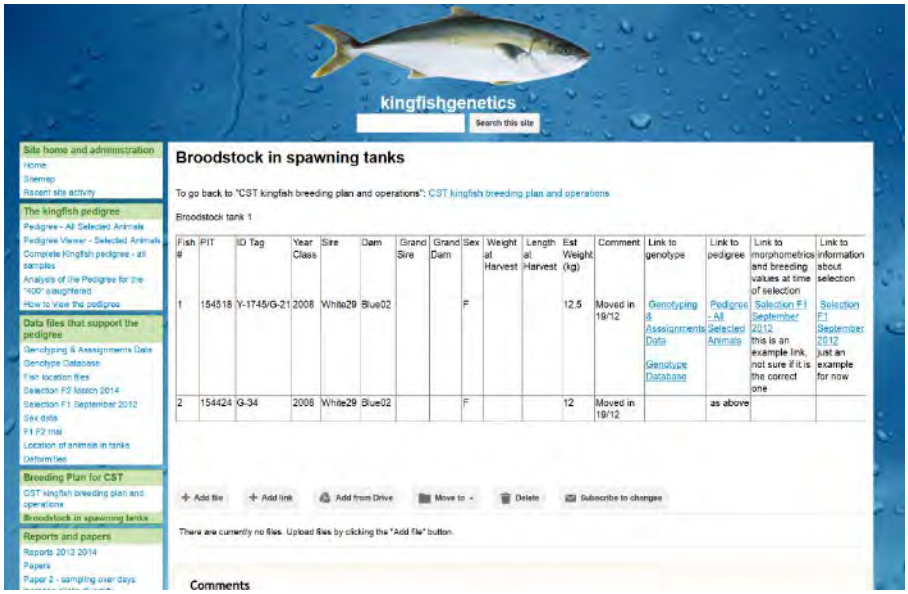
The web site also allows individual pages to be linked and present summarized data in a way that assists in a functional way to make breeding decisions.

For example, the following page gives a brief overview of the numbers of broodstock and tanks (FIGURE 3).

FIGURE 3. Numbers of animals in broodstock tanks

Clicking on “broodstock in spawning tank” link opens the following page (FIGURE 4), that contains detailed information about individual fish in the broodstock tanks.

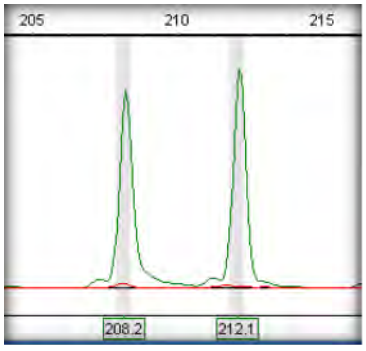
FIGURE 4. Details of individuals in broodstock tanks



From here (FIGURE 4) there are further links to:

- Genotyping and assignments data (containing all the genotypes and .fsa files listing individual animals; fsa files can be used to review the genotypes, even years after testing; FIGURE 5)

FIGURE 5. Electropherogram of genotype.

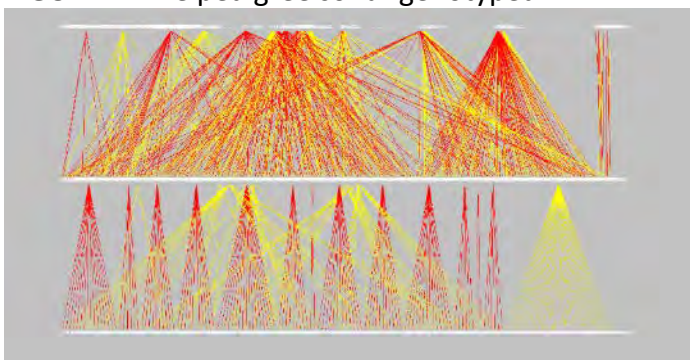


- The pedigree for all selected animals (FIGURE6 ; FIGURE 7).

FIGURE 6. e.g. pedigree records

Animal ID	Sire	Dam	Paternal Grandfather	Paternal Grandmother	Maternal Grandfather	Maternal Grandmother
$2013F_2$_102	$2008F_1$_115	Yellow18	Orange01	White28	Blue200	
$2010F_1$_54	$2008F_1$_172	Yellow249	Orange28	Yellow20	Orange01	$2010F_1$_54
$2013F_2$_68	$2010F_1$_40	$2008F_1$_125	Yellow250	Orange08	White29	Blue02

FIGURE 7. The pedigree so far genotyped



And finally,

- Information concerning the round of selection that produce that particular fish.

This latter web host hosts all the various files, photos, procedures and data from the particular round of selection producing the individual fish in question.

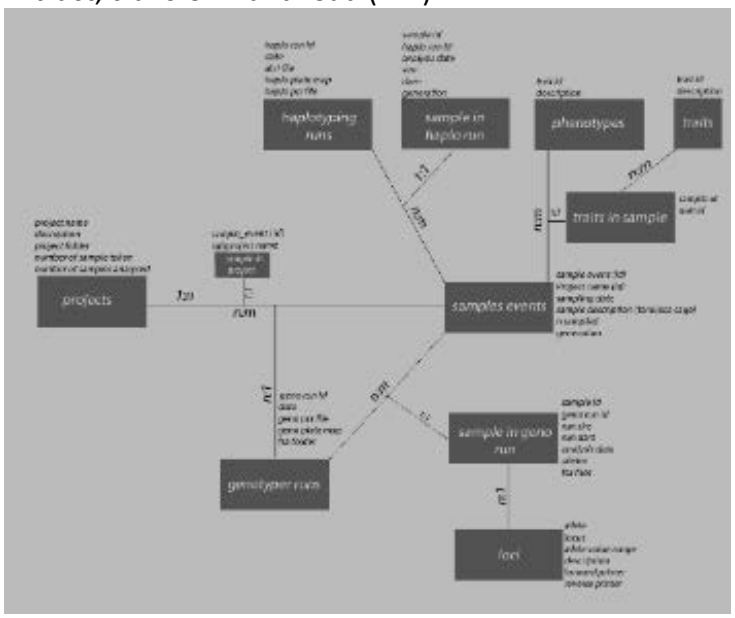
The decisive advantage of this web based system is that all staff can edit the site, add new items and features (including breeding values, ASREML analyses), and improve the functionality of the site. This aspect is not possible in more formal Access type data bases, usually requiring one or so specialist/ external consultants to design and keep up to date. The more formal data bases can create a critical dependency on the programmer and can cause time delays in decision making and management.

2.3 Highly structured data storage and retrieval

2.3.1 Designing a relational database

A data model was designed with an appropriate ERD diagram, (ERD= entity relationships diagram, http://en.m.wikipedia.org/wiki/Entity%E2%80%93relationship_model) linking the different aspects of the project (sampling, genotype, phenotype and pedigree data), into one relational database (FIGURE 8).

FIGURE 8.
Extract, transform and load (ETL)



We have imported data (genotype, phenotype, etc.) from the different Excel and Word files into an Access database. The data is solely stored in the Access database, where it cannot be changed by all users, but only authorized staff members under special circumstances such as mistypes or errors in the actual data. We then use MS PowePivot add-in (http://office.microsoft.com/en-au/excel-help/power-pivot-add-in-HA101811050.aspx) to pull the tables into an OLAP cube (http://en.wikipedia.org/wiki/OLAP_cube) and use an Excel-based user interface for any data investigation needed. As one of the predefined characters of the PowerPivot add-in, none of the analysed data could accidently be changed or lost through the user interface. In addition, the user interface allows researchers to gain direct reading access to both raw and analysed data. The OLAP cube engine in the background allows researchers to drill directly into the work done by any of the team members according to their needs.

2.3.2 Extracting and transforming the relevant data:

According to the data model relevant, data can be abstracted from sub projects, folders and files.

2.3.3 Data loading:

Once the data was uniformed it is uploaded it onto Ms Access.

2.3.4 OLAP Cube:

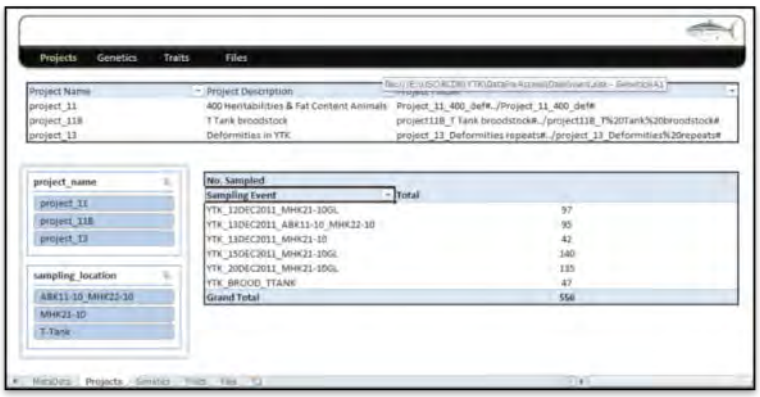
From the Access database, data are loaded into the PowerPivot window of the Excel PowerPivot add-in. Each Access table becomes one dimension of the relational database (as detailed below in the user interface section) and was then ready to be analysed.

2.3.5 Designing a user interface

We created an easy to use user interface, performing as a connecting layer between the user and the OLAP cube. Through the user interface the researchers may find their files and

folders as well as analyze and trace research data. The designed user interface contains various main pages, relevant to subprojects. All pages are connected through an easy access toolbar. The data on the different pages can be filtered and sliced. Slicers and filters are connected to the relevant tables to assure they are all sliced in the same time and thus avoiding mistakes.

FIGURE 9. User interface screen shot

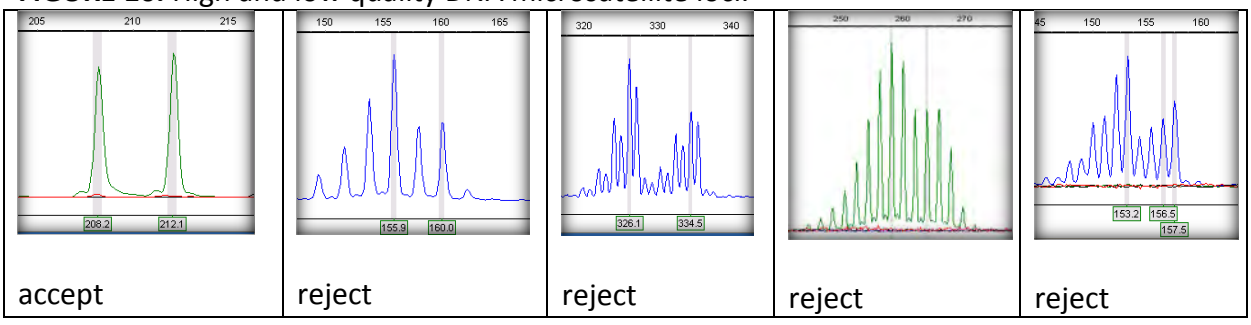


3 New microsatellite markers

We have synthesized and tested 60 primer pairs designed from GS-FLX 454 Next Generation Sequencing data obtained from two YTK transcriptomes (liver and muscle tissue). Thirteen of these are highly informative and are currently in use, together with four markers from published sources.

Examples of acceptable and unacceptable DNA microsatellite loci electropherograms (FIGURE 10) show that we restrict loci to those that are a. machine scorable (do not require human reinterpretation and thus errors), b. tend to not have any stutter, c. are polymorphic, d. amplify well, repeatedly and reliably, e. able to be panelled, given size and amplification characteristics. USC is the only group world-wide that have developed such a volume of high quality DNA microsatellite loci for kingfish.

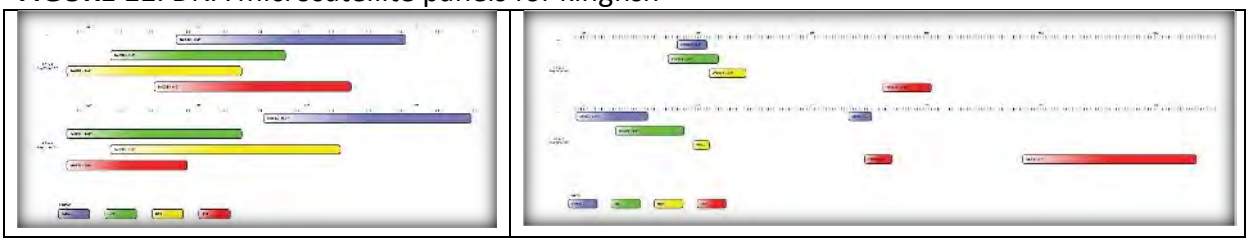
FIGURE 10. High and low quality DNA microsatellite loci.



accept	reject	reject	reject	reject

We have developed two PCR multiplex pools by which several microsatellite loci can be simultaneously amplified at reduced cost (FIGURE 11).

FIGURE 11. DNA microsatellite panels for kingfish



If more markers are still required, these can be developed from a pool of 500-800 potential SSRs remaining within the Next-Gen sequencing data.

Note: the ability to achieve high standard genotypes also relates the having a. high quality DNA extractions (USC has a "DNA robot" for rapid and uniform extractions), optimizing PCR cycles , and complete control of the operation and maintenance of an in-house genotyper (FIGURE 12).

FIGURE 12. Equipment used for high quality and high throughput genotyping.

DNA robot	PCR optimization	Genotyper
		

4 Estimation of $ value of genetic program including cost benefit analysis

Various data sets and various methods of analysis have been deployed to determine the commercial value of the CST genetic selection program. The best overall summary is that for every few cents CST invests into genetics, the genetic program returns about a dollar.

4.1 Spawning 29th October 2013

One day old larvae, spawned from "wild" and F_1 broodstock tanks, were stocked at 60 larvae per liter into 6 x 3000L larval tanks, with 3 replicates of each source of parental Kingfish, giving us 6 X 3000L tanks in total with larvae stocked at the same density (180k each tank).

Initial egg and 1 dph larval measurements show the F_1 eggs were 1.406mm diameter compared to 1.356mm diameter for the F_2. One-day old F_1 were 4.92mm long and the F_2 were 4.75mm long. So the F_1 eggs/larvae start around 3.5% bigger than the F_2's, which could be due to several factors such as fish age, maturity, genetics, and dietary intake.

Once hatched each broodstock source larvae were quantified and split into 3 replicate tanks

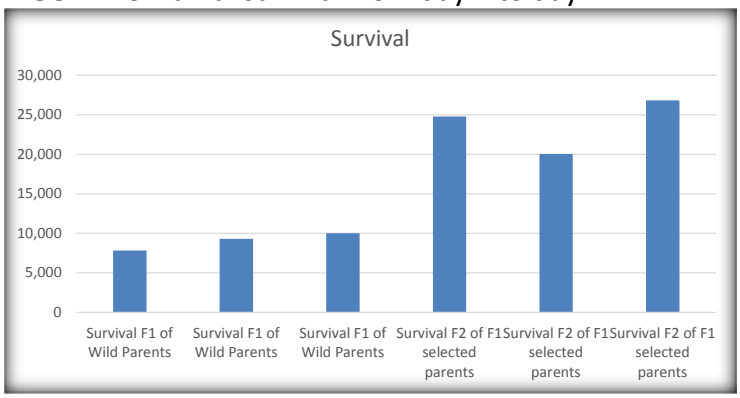
of each broodstock source.

Samples for DNA microsatellite analyses were taken at age 1 day.

4.2 Survival differences among F_1 and F_2 larvae from D1 to D21

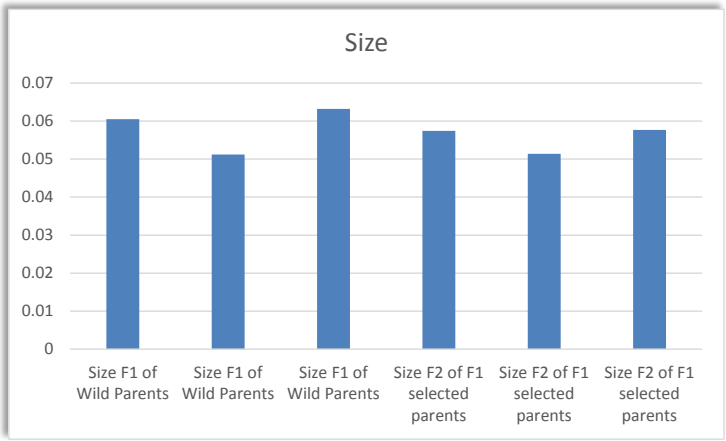
F_2 survival from day 1 to day 21 was repeatedly (in each of the three replicates) and up to several folds higher than that for the F_1 of wild broodstock (FIGURE 13).

FIGURE 13. Larval survival from day 1 to day 21



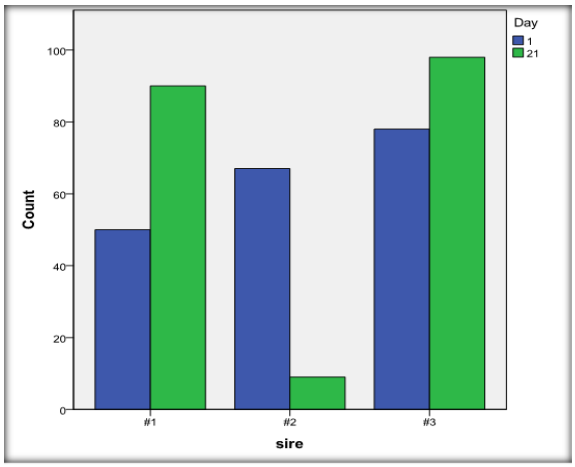
However, there could be non-genetic reasons for this difference, perhaps the eggs from the wild broodstock tank were poorly treated, or perhaps the F_2 fish were larger. The sizes of the F_2 larvae were not larger than that of the F_1s (FIGURE 14), therefor size differences at the start are not likely to account for the reported differences. Note: lower density should have promoted faster growth of the F_1 larvae over the F_2 larvae, other things being equal.

FIGURE 14. Sizes of the F_1 and F_2 larvae at 21 days post hatch.



Evidence that the survival difference could be genetic in origin could be deduced from observing that the offspring from specific wild sires (or dams) die off during hatchery rearing, but not those from the F_1 sires, and indeed this is the case. Specifically, offspring of wild sire #2 failed to thrive in the hatchery, and this observation was evident across multiple dams fertilized by sire #2 (FIGURE 15). No family based dropouts were observed for the F_2 fish. Note: 800 larvae were genotyped to obtain this information.

FIGURE 15. Differential survival of larvae from different sires



$$$$ from genetics: Value of increased larval survival †
A 3 fold increase in larval survival translates into more fish available per year for the same resources, or less existing infrastructure, staff and maintenance require to produce existing volumes.
Value = > 10^5

† assuming differences are genetic

4.2.1 Growth and FCR to day 84 on-land

Ex nursery fish (750 per origin) from each source were individually tagged and stocked into two 100m3 tanks (one tank per origin of fish).

The fish were fed to satiation 3 X daily and all feed has been recorded in order to measure FCR rates. Weight checks are taken every 3-4 days.

At 84 days post hatch, the F_2 fish averaged 91.14gr each with an FCR of 0.894. The F_1 averaged 79.64gr and had an FCR of 0.964, FCRs estimated since day 46 ph (TABLE 2).

TABLE 2. Growth and FCRs on F_1 and F_2 stocked in tanks

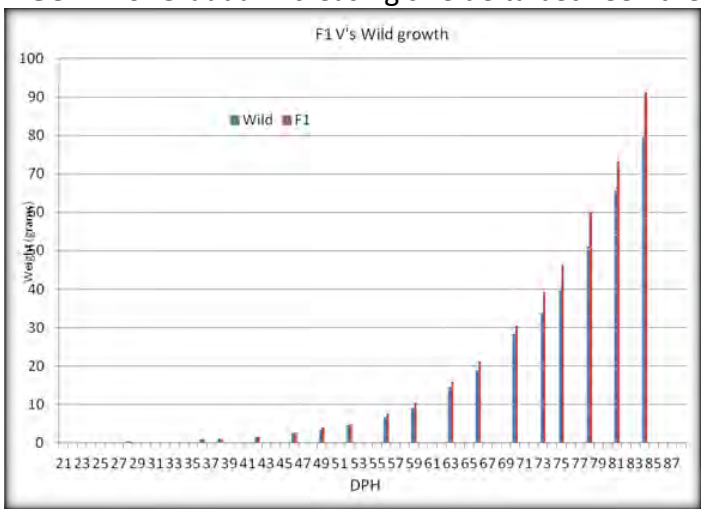
Since 13/12/2013	feed (gr)	biomass increase (gr)	FCR
Wild	54907	56940	0.964
F_1	58960	65972	0.894

The F_1 offspring has grown an extra 9032grams of biomass, but it has also eaten an extra 4053grams of food to get this growth.

Considering the growth trajectories over 87 days (FIGURE 16), it seems like there was a gradual but increasing size delta between the F_2 fish and the F_1 fish, despite the F_2 fish stating at a smaller size. However, these data should be considered as tentative as tanks

were not replicated.

FIGURE 16. Gradual increasing size delta between the F_1 and F_2 fish.



$$$$ from genetics: Value of increased growth and FCR †

In $ value, a simple calculation is the difference in $ gain from selling fish (9.032kg of weight gain x $18.5 per kg of fish = $167.1) and the expenses on extra food (4.053 kg of extra food x $2 per kg of feed = $8.1), i.e. 167.1 – 8.1 = $159 (at post-hatch with only 750 fish).

The economic benefit would be very large at harvest since genetic gain is also expressed in millions of fish under mass production.

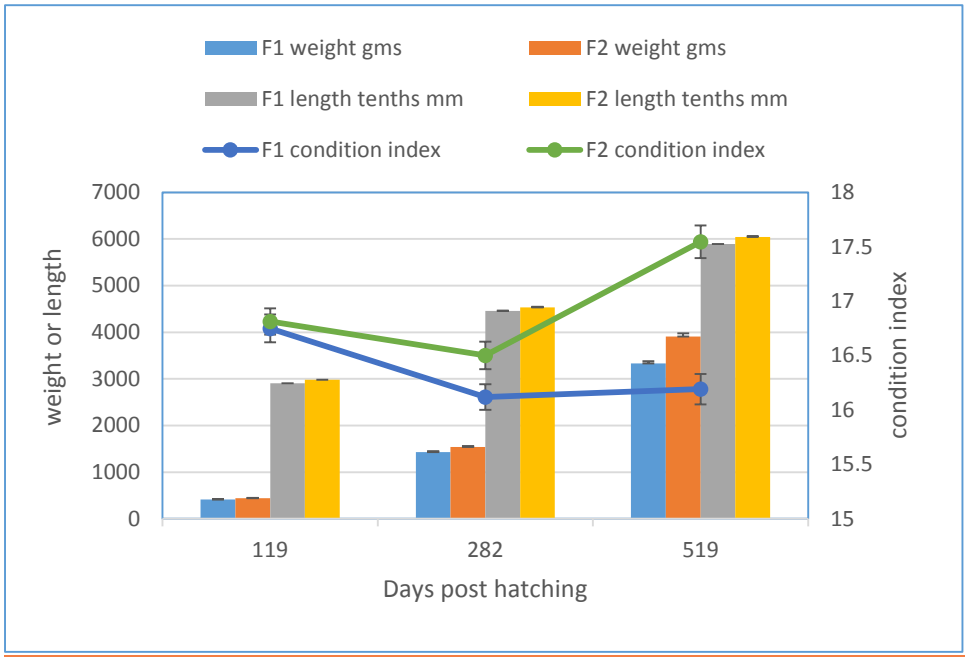
Value = > 10^6

† assuming differences are genetic

4.2.2 Communal rearing in single tank

At between 400 and 500g, F_1 and F_2 fish were tagged and put together into a single tank for communal rearing until they reached about 4 kg (FIGURE 17). Under these conditions, each single fish can be considered a replicate experiment, however any starting differences in fish weights and sizes need to be accounted.

FIGURE 17. Communal growth of F_1 and F2 fish from 0.4 to 4kg.



119 DPH corresponds to 25/4/2014, 282 DPH corresponds to 7/8/2014, 519 DPH corresponds to 1/4/2015

At the commencement of the communal rearing experiment in a single land based tank, the F_2 at day 119 DPH were already significantly heavier and longer than the F_1 (i.e. 7.492 % heavier, and 2.471% longer, ANOVA $F_{1,179} = 17.204$, $P < 0.001$; ANOVA $F_{1,179} = 16.790$, $P < 0.001$, respectively). The average condition index at 119 DHP was less than 1% different and was not significantly different between the lines.

By Day 163 DPH, the F_2 was 7.630% heavier and 1.689% longer than the F_1, so the relative advantage of the F_2 was little changed during this period, albeit weight was still significantly larger for the F_2 even after controlling for starting weights (ANCOVA, controlling for 119 DPH starting weights, $F_{1,176} = 5.650$, $P < 0.05$), but length was not significantly different between the F_1 and F_2 after accounting for the starting lengths. Between 119 and 282 DPH, condition index had increased considerably for the F_2 fish, relative to the F_1 fish, which now had a 2.374% higher index than the F_1, which was significantly larger even after controlling for the commencement condition indices (ANCOVA, controlling for 119 DPH condition indices, $F_{1,176} = 5.032$, $P < 0.05$).

At 519 DPH, the differences between the lines had accelerated, with a 17.264%, 2.600% and 8.360% superiority for the F_2 for weight, length and condition index, respectively. However, only the weight and condition index differences were statistically significant after controlling for the respective values at day 119 (ANCOVA, controlling for 119 DPH weight, $F_{1,168} = 27.280$, $P < 0.001$; ANCOVA, controlling for 119 DPH condition index, $F_{1,168} = 46.167$, $P < 0.001$; respectively). That is, after correcting for starting conditions, the F_2 fish grew much larger than the F_1s, not by increasing length so much, but by increasing condition as the fish recovered from winter conditions (during and after the 282 DPH data collection).

Given the growth advantage of the F_2s *after* controlling for starting weight, it is not unreasonable then to consider that at least some of the differences in starting weights were also genetic, especially at the beginning of the separate rearing trial, when the F_2s were slightly smaller than the F_1s. Therefore removing these starting weight differences is evident at the end of the separate rearing trial and would cause the final calculation of genetic differences to be underestimated.

$$$$ from genetics: Value of increased growth to 4kg †

Previously, we predicted what would be the selection response for weight by the following equation:

$$R = h^2 S$$

Where

- *R* is the selection response for fish from a cohort at average harvest weight of 3.0725 kg.
- h^2 is the heritability for weight from a cohort at average weight 3.0725 kg, as is assumed to be 0.30.
- *S* is the selection differential, and was the average weight of the selected animals, 3.7532 kg, minus the average of the random unselected animals, 3.0725, which is equal to 0.6807kg.

Thus, *R* was calculated to be 0.20421 kg, ie 204 gms per fish in terms of genetic gain, which is worth about $2.04 cents per fish (assuming a fish is worth $10 per kg), prior to subtracting feed costs. If we use fish that average 4kg after selection, which may be possible, the added value per fish is $2.783.

For the present tank trial, we find about a little over 575g difference between the F1 and F2 at 4kg. So the empirical testing of F2 vs F1 is about twice that predicted from the original rounds of selection, with perhaps the unanticipated bonus due to domestication selection.

Value = >> 10^6

†assuming differences are genetic

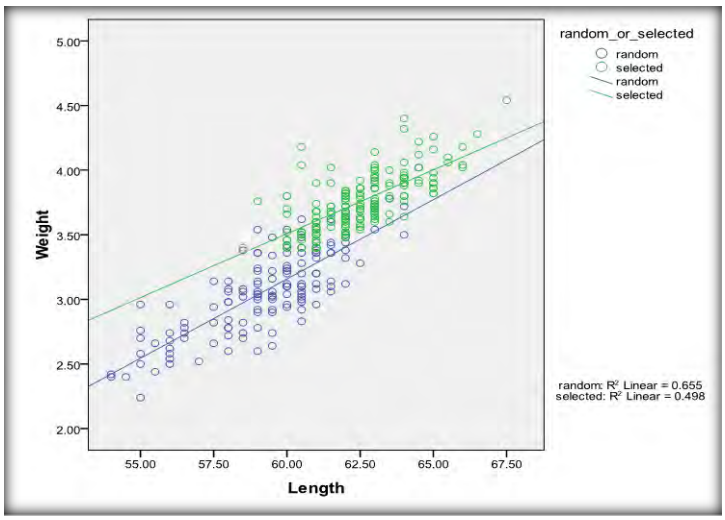
4.3 Other metrics (other than growth and survival) that indicate a commercial return from the genetic program

Were there any other metrics considered that could also point to genetic changes between the F_2 and F_1?

4.3.1 Change in body proportions and condition index

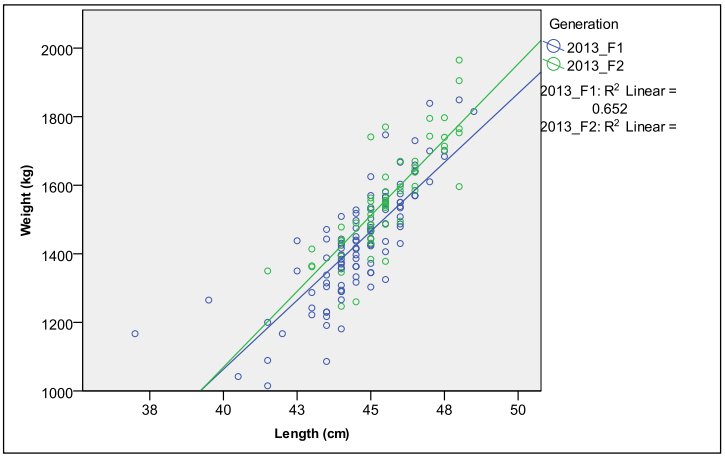
In previous reports we showed that the selection for the F_1 fish was based not only on selection for absolute size, but also on fish being “heavy for weight” and Figure 18 gives an example of the distributions of the unselected F1 and selected F_1 from one of the prior selection experiments.

FIGURE 18. Regressions of weight on length for the selected and random fish (from the round of selection that produced the F_1 parents)



We can now look at the unselected F_2s and F_1s in the present tank trial: In the next generation, the progeny of the selected fish, the F_2s, tend to be heavier for length than the F_1s (FIGURE 19), and indeed the condition index was subtaintially higher for the F_2 and the F_1 fish.

FIGURE 19. Regressions of weight on length for the F_1 and F_2 fish.



4.3.2 Pictorial evidence for a change in body shape

Fish were selected not only for large total weight, but selected for high weight to length ratios, and this is reflected in the changed morphology of wild, to F_1 and to F_2.

FIGURE 20. Wild fish tend to be relatively long and skinny.

FIGURE 21. Selected F_1 fish have deeper bodies.

FIGURE 22. F_2 (unselected) fish appear to have inherited the deeper body type.

$$$$ from genetics: Value of increased condition index †
A better condition index makes the fish eligible to be sold in Japan.
Value = > 10^5

† assuming differences are genetic

4.3.3 Industry observations: biomass increase

Since using the genetically selected stock for production, CST records show the kingfish biomass (monthly) and cumulative (FIGURE 23; FIGURE 24; from CST's annual report) greatly exceeds that forecast for regular genetically unimproved stock.

FIGURE 23. Monthly kingfish biomass increase

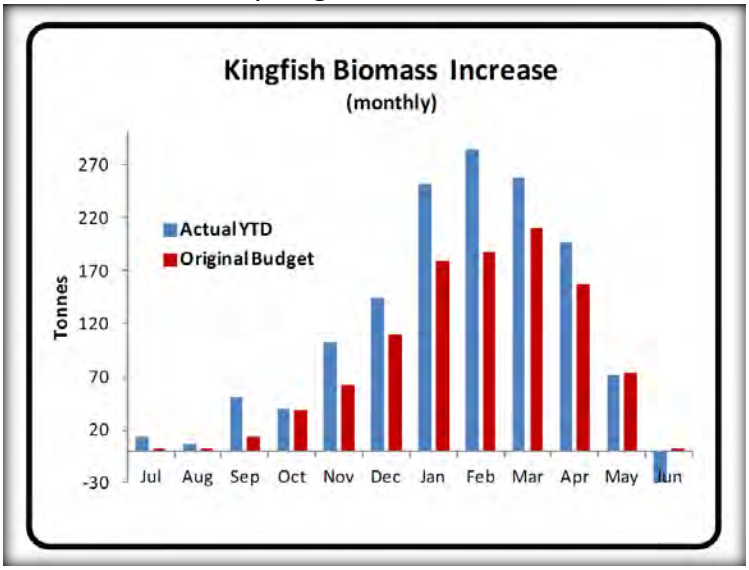


FIGURE 24. Cumulative kingfish biomass increase

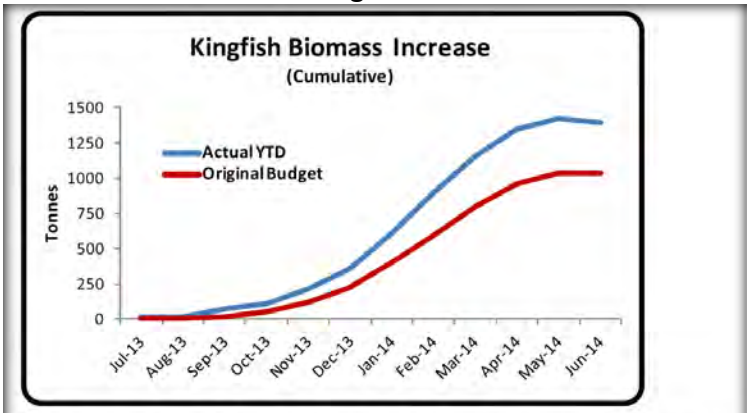


Figure 2 shows our actual biomass increase (blue) compared to our budget biomass (red).

4.3.4 Anecdotal reports from Clean Seas regarding growth of various F_1 and F_2 sea cages in recent times

- the F_2's survival rates are as high as ever and growth rates are above expectations, even though the water temperature is 1.5^0C lower than average at this time of year.
- The (F_2) fish are growing big so quickly it is putting severe pressure on the harvesting of the fish, as the cages are becoming dangerously high with biomass stocking.

4.3.5 CST becomes profitable during the first year it used genetically improved stock

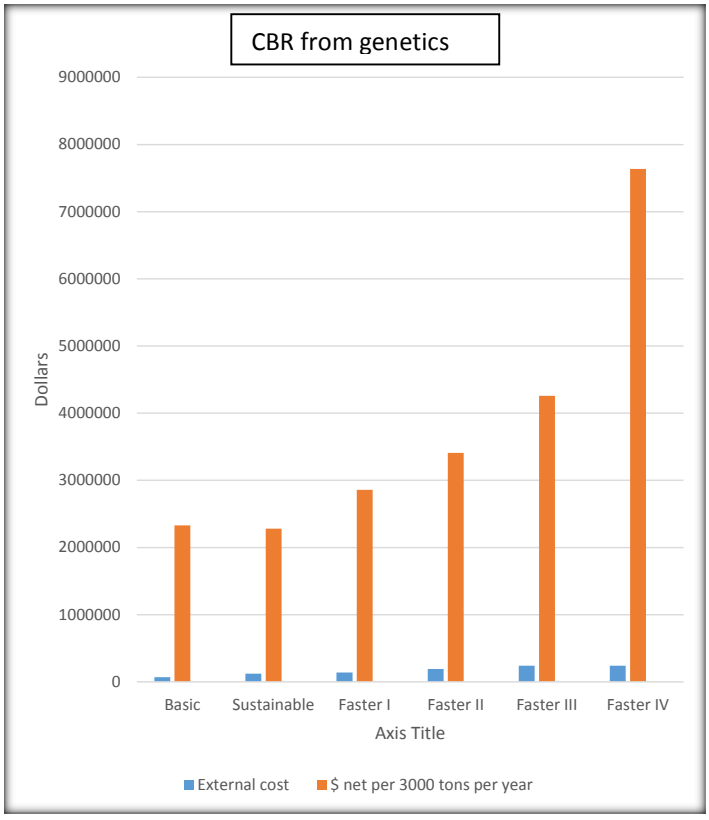
CST, a public company moved from reporting losses to profit during the first year the genetically improved stock were. This is the second time we have observed such a coincidence of profitability for a given company with the full deployment of genetically improved stocks.

4.4 Cost benefit returns

Various economic and genetic models and assumptions were considered. Economic return from genetic gain set at between 2 to 7.5%, the crop was estimated to be 3,000 tons,

returns were estimated at 2018, 2022, 2026, investment costs considered extra external costs (mostly genotyping). As usual for genetic programs (eg. Pig in UK, salmon in Norway) very impressive CBRs are predicted. The models did not indicate a plateauing of return from investment (FIGURE 25) suggesting, all other things being equal (and they are not), there was a business failure or underinvestment at the lower range of investment (basic), or looking at it the other way, an unexploited business opportunity exists for CST. For example, at the lower range of investment, about 5 cents in every $ earned by the genetic program is returned to the program for external genetic support. Doubling that to 10 cents would, theoretically, return far more than the cost of the investment. It was reported at WAS, Adelaide, that the NZ King salmon genetic program return 33 cents in the dollar earned by their genetic program back into genetics R&D. Note: the most recent results of nearly 20% genetic gain suggest the preceding estimates could be underestimates.

FIGURE 25. CBR from genetics



4.5 Conclusions on $ commercial value

Slowly, and over time, the fruits and commercial benefit of the planning and implementation of a sustainable YTK genetic program are being realized, with incoming scientific data showing that selection responses are up to about 20% per generation. Economic modelling of this benefit indicate cost benefit returns typically greater than 10:1.

Presently, the CRC partnership and investment have allowed CST to become the only functional, commerical and sustainable kingfish genetic program in the world, which could

bring substantial market advantages to the company. While others have attempted some type of "genetic" work, they have failed to develop a cross generation genetically sustainable breeding nucleus, typically failing even at the first generation (the WA pedigree has only one female and one common male, and offspring cannot be used to form a breeding pedigree due to inbreeding). The needs, infrastructure and the detailed planning to maintain genetic diversity of say 40 families at the F_1 and F_2 are much much greater than those at G_0, and these planning and infrastructure challenges have been solved **uniquely** at CST and the genetic operations are now embedded in the industrial production cycle. How long these conditions of competitive advantage will endure for CST is unknown. On one hand, the logistics, infrastructure, trained staff (in genetics), experience, specialist hatchery and statistical skills that have been drawn together and that opearate now as a functional unit are quite formidable, and it is unlikely rival teams could be compiled and competitive in the near future. Probably, at least within the next several years, a greater threat to CST is more likely the loss of its genetics MO, refined operating procedures, teams and their present functionality and integration, rather than threats from rivals.

Soon the F_3s should be available for production, and with them a corresponding commerical gain.

5 Plan for the next five years of the genetic program and breeding objectives reviewed

A meeting was held between CST and USC in late 2014. The breeding program for the future was developed as well as breeding objectives. The following executive summary outlines these items.

5.1 Executive summary

Company breeding objectives, set by management, are to focus the genetic program in directions that best support the economic performance of the company. The set objectives include improving time to market of big, well conformed fish, reducing generation time and maximizing genetic diversity.

Methods to technically deliver the breeding objectives are optioned and focus on selection at sea for three traits (growth, condition index and appearance). Practical methods of keeping maximum genetic diversity are scoped.

An inventory is catalogued for CST genetic assets, which presently includes a legacy of selected fish, corporate knowledge in genetic selection that is well documented in reports, infrastructure for broodstock and staff expertise.

6 Shortening generation time

A key to increasing profitability of genetic programs is the shortening of the generation time. To address this, we have produced recombinant YTK follicle stimulating hormone (FSH) and used it in *in-vitro* and *in-vivo* trials.

6.1 Methods

6.1.1 Production of Recombinant Yellowtail Kingfish Follicle-Stimulating Hormone (recYTK_FSH) in *P pastoris*

Molecular cloning of FSH β and α subunits and design of expression construct

Total RNA was extracted from pituitaries of 3-year old YTK (CleanSeas), and first-strand 3' and 5' cDNA were synthesized according to Nocillado et al. (2013). The full-length FSHβ and FSHα cDNA sequences were isolated using Rapid Amplification of cDNA Ends polymerase chain reaction (RACE PCR). The gene-specific 3' and 5' RACE PCR primers were designed from our previously isolated partial YTK FSHβ sequence (HQ449731). The 3' RACE PCR primers were 5'CCATATGTGCAGGACAGTGTTACC3' and 5'GAGGTGAAACACACTCCTGGATGT3' (first and nested, respectively). The 5' RACE PCR primers were 5'ACATCCAGGAGTGTGTTTCACCTC3' and 5'GGTAACACTGTCCTGCACATATGG3' (first and nested, respectively). The full coding sequence of YTK FSHα was isolated by 3'RACE PCR using as primers 5'GGAAACTTTCTCTCAATGTGGTGAC3' and 5'GCTGCAACCACGATGGGCTCAGTG3'. These primers were designed according to the conserved sequences of the European seabass (Accession no. AF2691457) and chub mackerel (Accession no. JF495131) FSHα sequences. Products were cloned as described in (Nocillado et al. 2013) and sequenced at the Australian Genome Research Facility. Sequences were analysed and assembled with Sequencher Version (Genes Codes). The signal sequence of the FSH α and β subunits were identified using SignalP 4.1 software (Petersen et al. 2011) and were not included in generating the expression construct.

A single chain expression construct was designed according to Aizen et al (2007) and illustrated in Figure 26. The FSH α and β subunits were linked with a linker sequence (GSGSHHHHHHGSGS) that contains 6 histidine (6His) tag. The standard coding sequence was then codon-optimised according to the yeast (*Pichia pastoris*) coding preference and then the construct synthesized by a commercial provider (GenScript USA Inc., Piscataway, NJ). GenScript generated the expression construct by directional ligation of the synthesized single chain FSH sequence into the *EcoR*I-*Not*I restriction sites (that were also added on the respective ends of the FSH sequence) of the yeast expression vector pPIC9K. The insert sequence and orientation were verified by sequencing (GenScript).

The recYTK_FSH expression construct was used to transform competent DH5α *E coli* cells. Positive colonies were identified by PCR and grown in liquid nutrient media. Plasmid DNA was purified using Qiagen's miniprep purification columns.

6.1.2 Production of recYTK_FSH in yeast (P pastoris)

In preparation for electroporation into *P pastoris*, the plasmid DNA was digested with *Sal*I (New England Biolabs) according to the manufacturer's protocol. The digested DNA was purified through miniprep purification columns (Qiagen). Circular pPIC9K vector DNA, which was used as negative control for electroporation, was similarly prepared. Eight micrograms of purified *Sal*I-digested construct was used to electroporate GS115 Superman5

His- strain of *P pastoris*. Positive transformants (histidine positive) were selected based on survival in histidine-deficient media. Dilution series in YPD media containing agar, protease peptone and yeast extract were used to screen for high copy number colonies based on the resistance to Geneticin® (G418 sulfate, Invitrogen). Protein expression induction was carried out in a shaker flask at 30˚C. Yeast was harvested 72h after induction by methanol. The recYTK_FSH was purified using magnetic nickel nitrilotriacetic acid-agarose beads (Ni-NTA; Qiagen) that capture 6His-tagged proteins. All of the mentioned steps including electroporation protocol, selection in histidine-deficient media, Geneticin resistance, and the protein expression induction by methanol were carried out according to Invitrogen's *P. pastoris* expression manual (Invitrogen) and Aizen et al. (2007).

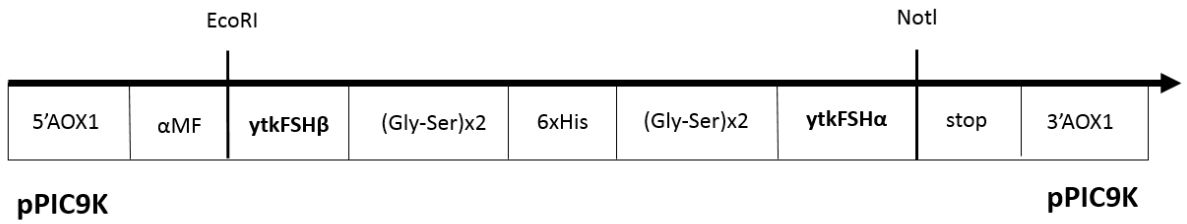


FIGURE 26. Diagram of the recYTK_FSH construct.

Production of recYTK_FSH *in yeast* (*P pastoris*)

In preparation for electroporation into *P pastoris*, the plasmid DNA was digested with *Sal*I (New England Biolabs) according to the manufacturer's protocol. The digested DNA was purified through miniprep purification columns (Qiagen). Circular pPIC9K vector DNA, which was used as negative control for electroporation, was similarly prepared. Eight micrograms of purified *Sal*I-digested construct was used to electroporate GS115 Superman5 His- strain of *P pastoris*. Positive transformants (histidine positive) were selected based on survival in histidine-deficient media. Dilution series in YPD media containing agar, protease peptone and yeast extract were used to screen for high copy number colonies based on the resistance to Geneticin® (G418 sulfate, Invitrogen). Protein expression induction was carried out in a shaker flask at 30˚C. Yeast was harvested 72h after induction by methanol. The recYTK_FSH was purified using magnetic nickel nitrilotriacetic acid-agarose beads (Ni-NTA; Qiagen) that capture 6His-tagged proteins. All of the mentioned steps including electroporation protocol, selection in histidine-deficient media, Geneticin resistance, and the protein expression induction by methanol were carried out according to Invitrogen's *P. pastoris* expression manual (Invitrogen) and Aizen et al. (2007).

Analysis by Western blot

The recombinant protein was further purified by dialysis in Slide-A-Lyzer G2 Dialysis Cassettes (Thermo Scientific) according to the manufacturer's protocols. The dialysis

cassettes eliminated peptides with molecular weights less than 20 KDa. The estimated molecular weight of recYTK_FSH was 23 KDa based on its amino acid sequence. Deglycosylated recYTK_FSH were analysed by Western blot. Deglycosylation was performed by incubation at 37°C of heat-denatured 20 µg recombinant protein in *N*-glycosidase F (PNGase F) based on supplier recommendations (New England Biolabs) in a total reaction volume of 25 µl. A reaction without PNGase F was performed as negative control. Twenty-five microliters of the reaction was electrophoresed on 12% polyacrylamide gel (BioRad Mini-Protean TGX). Gel was blotted onto nitrocellulose membrane according to BioRad's Transblot Turbo System. The membrane was blocked with 3% BSA prior to incubation with the anti-His primary antibody according to the manufacturer's (Qiagen) protocol. The washed membrane was incubated in goat anti-mouse secondary antibody conjugated to an infrared dye (Licor). After another buffer wash, the membrane was viewed on an infrared imaging system (Odyssey CLX, Licor).

Determination of the bioactivity of recYTK_FSH

Luciferase reporter gene assay

A reporter gene assay according to Biran et al (2008) was conducted in order to determine whether the recYTK_FSH can bind and activate the FSH receptor. Briefly, the reporter gene used was luciferase, which is transcriptionally regulated by cyclic AMP (cAMP) response element (CRE; Invitrogen). Activation by recYTK_FSH was tested on the tilapia FSH receptor. The assay was conducted at the laboratory of Prof. Berta Levavi-Sivan at The Hebrew University of Jerusalem.

In vitro bioassay

In order to confirm its capacity to stimulate the secretion of sex steroids, *in vitro* bioassays of the recYTK_FSH were conducted as described in Aizen et al. (2007) in tilapia and in YTK gonadal fragments. Tilapia ovarian fragments (20 mg) were obtained from mature female (GSI 2.66%) fish. The fragments were washed three times by incubating them with gentle shaking in 96-well plates for 3 hours, with the culture media being replaced every hour, at 28°C. The culture media was basal medium eagle (BME; Sigma) that contained 2.2 mM $NaHCO_3$, penicillin (50 IU/ml), streptomycin (0.05 mg/ml), nystatin (1.25 IU/ml), and 0.05% BSA (Sigma). The pH was adjusted to 7.4 with 2.1 mM of Hepes buffer. After the third wash, the fragments were incubated in fresh culture media (1 ml/well) that contained recYTK_FSH at concentrations of 125 ng/ml, 250 ng/ml and 500 ng/ml. Fragments incubated in media only comprised the negative controls. Positive controls were those incubated in culture media containing tilapia pituitary extracts, which were obtained from mature female tilapia. Undiluted and 10-fold dilutions of pituitary extracts were used. The incubations were performed in triplicate per treatment group. After 18 h incubation, the media were then collected and stored at -80°C until extraction of steroids and E_2 immunoassay.

For the *in vitro* bioassay in YTK, ovarian and testicular fragments (20 mg each) were used for the experiment. Gonadal fragments were obtained from prepubertal YTK (>1 yr old; mean body weight) and washed as described in the tilapia experiment. Then, the fragments were incubated in culture media (as in the tilapia experiment), however the recYTK_FSH doses

were 800 pg/ml, 400 pg /ml, 200 pg /ml and 100 pg/ml. The doses used were lower than those used for the tilapia experiment because the recombinant peptide is homologous. Incubation temperature was at 21°C. The duration of incubation and media collection and storage was as described for tilapia.

Steroids from the culture media were extracted with diethyl ether. Levels of E_2 and 11KT were determined using commercially available specific ELISA kits (Cayman Chemical, Ann Arbor).

In vivo trial in YTK

In vivo experiments at CleanSeas Tuna Ltd (CST, Arno Bay, South Australia) commenced on the 9th of October 2014 and were terminated on 9th of December 2014 (9 weeks). The experimental animals were kindly provided by CleanSeas. Previously tagged fish were randomly assigned into 3 experimental groups, n=15/group. At the start of the trial, the gender of the fish could not be determined. Lyophilised recYTK_FSH was resuspended in sterile saline (0.9% NaCl) for injections at doses of 10 µg/kg and 20 µg/kg. Controls were injected with saline only. Injections were repeated once every 10 days for 9 weeks. During the experiment, the fish were fed daily with commercial feeds according to standard YTK feeding protocol at CleanSeas.

At the end of the trial, total weight, gonad weight and fork length (FL) were recorded. Gonadosomatic index (GSI) was calculated ((gonad weight/body weight) x 100). In addition, fish were bled, and brain, pituitary, gonad and liver were collected and stabilised in RNALater (Ambion). Plasma from blood was collected by centrifugation at 4,000 rpm for 20 min at 4°C. Plasma was stored at -80°C prior to analysis for E_2 and 11KT levels using specific ELISAs (Cayman Chemical, Ann Arbor).

For gonadal histology, samples from the right and left gonads were removed from each fish (n=43) and fixed in 4% paraformaldehyde for 24h and then transferred to 70% ethanol. Cross sections were taken from the medial area of both gonads, embedded in paraffin and sectioned transversely at 6 µm. Sections were stained with haematoxylin and eosin. Oocyte diameters (n=20/fish) were measured from left and right gonad of each female. Maturational stages were determined according to Poortenaar et al. (2001).

6.1.3 Oral delivery of peptides and DNA construct

Liposome-encapsulated GnRHa

Liposome nanoparticles (100 nm) containing GnRHa were prepared in a mixture of phosphatidylcholine, cholesterol and chitosan complex according to a protocol optimised at the laboratory of Prof Ming-Wei Lu, Aquaculture Department, National Taiwan Ocean University. GnRHa content in the liposomes was determined by a GnRHa-specific ELISA (Mylonas et al. 2007).

Chitosan-encapsulated GnRHa

Encapsulation of GnRHa in chitosan-alginate nanoparticles was performed according to Fonte et al. (2012). The amount of GnRHa in the nanoparticles, and supernatant during nanoparticle preparation, was determined by HPLC and used as an indicator of the efficiency of encapsulation. A GnRHa standard curve was generated first by HPLC according to Kafka et al. (2010) with minor modification. Briefly, reverse phase HPLC was carried out using a system consisting of a Perkin Elmer series 200 pumps and autosampler, a Fusion-RP column (C18, 75 x 4.60 mm i.d., 4µm particle size, 80Å, Phenomenex, CA) and a Flexar PDA wavelenght detector. The injection volume for all samples was 50 µL. The mobile phase consisted of solvent A: 0.1% (v/v) acetic acid in milli-Q water and solvent B: CH_3CN containing 0.1% (v/v) acetic acid. To equilibrate the system, solvent B at a concentration of 5% in solvernt A was run over 1 min at a flow rate of 1.2 mL/min. Samples were eluted by a gradient development from 5 to 40% of solvent B in solvent A at a flow rate of 1.2 mL/min over 12 min. To re-equilibrate the system, the concentration of solvent B was decreased to 5% over a period of 3 min. The detection wavelength was 220 nm.

The GnRHa nanoparticles were freeze-dried and stored at -20°C. Feeds containing GnRHa feeds were prepared by mixing the dried GnRHa nanoparticles with fish meal, potato strarch and fish oil at a ratio of 5:2:1 by weight. The GnRHa fish feeds were stored at -20°C until use.

For feeding trial (this is an ongoing study), 3 groups of tilapia (3 females and 2 males per group) were randomly assigned to experimental tanks. The negative control group are fed with blank pellets. The positive control fish were injected with GnRHa at a dose of 50 µg/kg and fed with blank pellets. The group treated with chitosan- encapsulated GnRHa were fed with pellets containing the nanoparticles at a dose range of 300-600 µg/kg. The water temperature is being maintained at 26-28°C. Fish are monitored for mating and spawning behaviour.

Recombinant Southern Bluefin tuna gonadotropin-releasing hormone 1

The full coding sequence of Southern Bluefin tuna (SBT) gonadotropin-releasing hormone 1 (sbtGnRH1) was isolated by 3'RACE PCR. First-strand 3'RACE cDNA was synthesised from SBT brain total RNA. The 3'RACE PCR primers were designed according to the Northern Bluefin tuna GnRH1 sequence (Accession no. EU 239500). The first round primer was 5'CAAATCAGAGAAGCAGCTTGCC3' while the nested primer was 5'ATGCACAGAAGAATGGCTATG. The cloning method, sequencing and sequence analysis was as described in Section 2.1. The region encoding for the decapeptide together with the GnRH-associated peptide (GAP) was amplified with added *BamH*I and *Not*I restriction sites at the 5' and 3' ends, respectively. In addition, the yeast consensus Kozak sequence (5'CATAATGTCT5') was added before the cleavage site (KR), which is immediately flanking the 5' end of the GnRH1 decapeptide. The construct diagram is shown in Figure 27. The construct was then ligated into the yeast expression vector pPIC3.5K, which does not contain a signal sequence thereby preventing the release of the expressed peptide into the culture media. The preparation of the plasmid construct for introduction into yeast, electroporation protocol into *P pastoris*, transformant selection and culture of positive transformants was as described in Section 2.2. Expression of the recombinant peptide was

verified from yeast protein extract by MALDI TOF/TOF mass spectrometry.

FIGURE 27. Diagram of the recombinant Southern Bluefin tuna GnRH1 construct.

Recombinant Southern Bluefin tuna luteinizing hormone

The sequence encoding for Southern Bluefin tuna luteinizing hormone β subunit (SBT LHβ) was isolated by 3'RACE PCR, using as template first-strand 3'RACE cDNA, which was synthesised from SBT brain total RNA. The 3'RACE PCR primers were designed from the Northern Bluefin tuna (Accession no. EF205591) LHβ sequence. The first round and nested 3'RACE PCR primers were: 5'CGCGGGGAGCACACCGGCGACAGAC3' and 3'CCAGAAAGGACGATGGCTGTACAAG3' (respectively). The PCR protocol, sequencing and sequence analysis were as described in Section 2.1. The SBT LHα subunit was also isolated by 3'RACE PCR as in the LHβ subunit. The 3'RACE PCR primers were the same as those used for the isolation of the YTK FSHα subunit described in Section 2.1.

Two single chain constructs, one linked with glycine (G) and the other with the carboxy terminal peptide (CTP) of the human chorionic gonadotropin, was generated by overlapping PCR as described in Moles et al. (2011), with modifications depending on the compatible restriction sites. The fragments linked are shown in the diagram in Figure 28. The sequences of the overlapping primers are shown in Table 2.

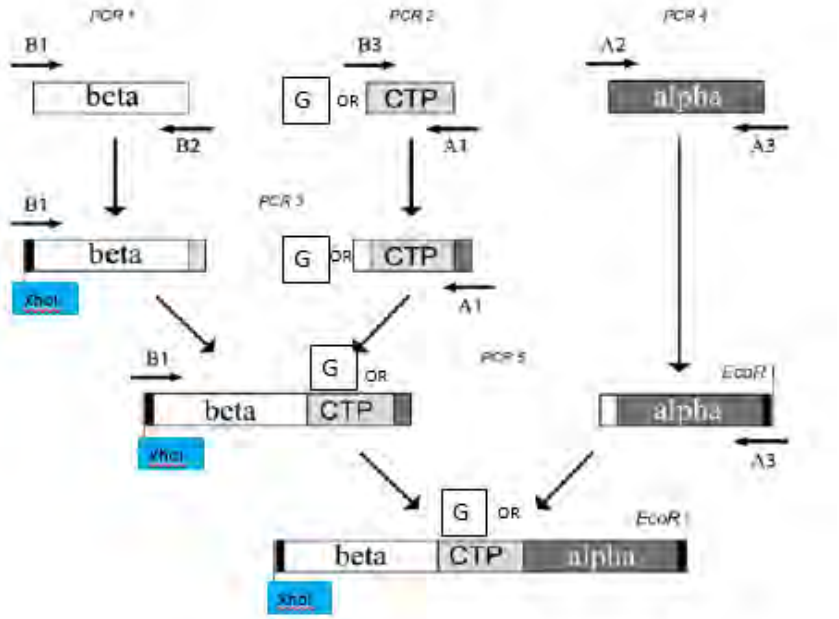


FIGURE 28. Diagram of the recombinant single-chain Southern Bluefin tuna LH construct generated by overlapping PCR. The LH β and α subunits were linked either with a glycine (G) or the carboxy terminal peptide (CTP) of the human chorionic gonadotropin. The corresponding primers for amplification are shown in Table 7.

Table 7. Sequences of primers used to generate the single chain Southern Bluefin tuna LH construct by overlapping PCR.

Primer code	**5' to 3' sequence**
B1	TACTCTCGAGCCAGAAAGGACGATGGCTGTAC
B2 (CTP linker)	TGAGGAAGAGGAGTAGTAGAAAGGTATGTC
B3	CCTTTCTACTACTCCTCTTCCTAAAGGC
A1	GAACATTGGGGTATTGTGGGAGGATCGG
A2 (CTP linker)	ATCCTCCCACAATACCCCAATGTTGACTTA
A3	CAGAATTCGGGTGCACATTTTAATTGC
B2 (G linker)	CAACATTGGGGTAGCCGTAGTAGAAAGGTATGTC
A2 (G linker)	GGCTACCCCAATGTTGACTTATCAAACATG

6.1.4 Transcriptomics: Yellowtail Kingfish Liver

Total RNA was extracted from liver of 6-8 kg YTK (2-3% GSI; CleanSeas) using Direct-zol RNA miniprep kit (Zymo Research) that includes a genomic DNA elimination step. The total RNA concentration was measured by a fluorometer (Quantus, Promega). RNA integrity was verified on the BioAnalyzer (Agilent). The absence of contaminating genomic DNA was confirmed by performing a PCR using housekeeping gene (acidic ribosomal phosphoprotein 0) and an aliquot of the total RNA as template. One microgram of total RNA was used to construct a transcriptomic library (separate library for male and female; each had a specific barcode) according to Illumina's TruSeq RNA sample preparation v2 guide, low sample protocol (Illumina).

Libraries were sequenced using illumina MiSeq benchtop sequencer at the University of the Sunshine Coast producing 250 bp paired-end (PE) reads. Fastq files were recovered from the sequencer and quality of the raw sequencing reads were assessed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimming of adaptors, filtering of low quality reads and removing contaminants was performed using Trimmomatic (Bolger et al. 2014) with a custom screening database. Clean PE reads from the MiSeq platform were assembled using Trinity (release r20140413) (Haas et al. 2013), using the default settings for paired-end reads except for: --min_contig_length 100, --normalize_reads. The assembled transcripts were screened for putative protein-coding regions using the program TransDecoder (release r20140704) (http://transdecoder.github.io/) with the option of including matches to the Pfam domain database (Finn et al. 2014), and with a minimum size of 30 amino acids (aa), then clustered using CD HIT with a sequence identity threshold of 0.9. Protein sequences were aligned against the NCBI KOG database (version 2/2/2011) using RPSBLAST 2.2.15 using an E value cut off of 1e-5 and performed via the WebMGA server (http://weizhonglab.ucsd.edu/metagenomic-analysis/) (Wu et al. 2011). Sequence similarity searches were performed using PFam hit data and BLAST searches.

6.2 Results and Discussion

6.2.1 Kisspeptin *in vivo* experiment

In order to achieve the objective of shortening the generation time of YTK that have been genetically selected, two approaches were undertaken. One was induction of sexual maturation using the kisspeptin peptides. We have previously shown that the kisspeptin decapeptides can stimulate the onset of gonadal maturation in immature male YTK (Nocillado et al. 2013). We also have previous data showing that kisspeptin-2 decapeptide can stimulate oocyte growth in immature female YTK (Nocillado et al. 2012).

In the present study, we utilised the conserved 15 amino acid kisspeptin-1 peptide (Kiss1-5) and the 12 amino acid kisspeptin-2 (Kiss2-12) peptide to advance puberty onset in immature female YTK. Based on plasma E_2 levels, neither Kiss1-15 nor Kiss2-12 exerted a positive effect (Figure 29). Positive response was not also observed in the positive control (GnRHa treatment). Due to limited sampling accessibility, we could not perform histological analysis of the gonads. We have yet to develop an ELISA for YTK FSH, which would enable to indicate whether the treatment had an effect at the pituitary level. We suspect the dose (50 µg/kg vs 100 µg/kg in our previous work) and frequency (every 3 weeks vs fortnightly in our previous experiments) of treatment might have contributed to the lack of significant response in the present study. Nevertheless, with longer duration of the treatment, a definite response might have been observed. The somewhat increasing mean E_2 level in the group treated with Kiss1-15 suggested advancing gonadal development as the treatment progressed.

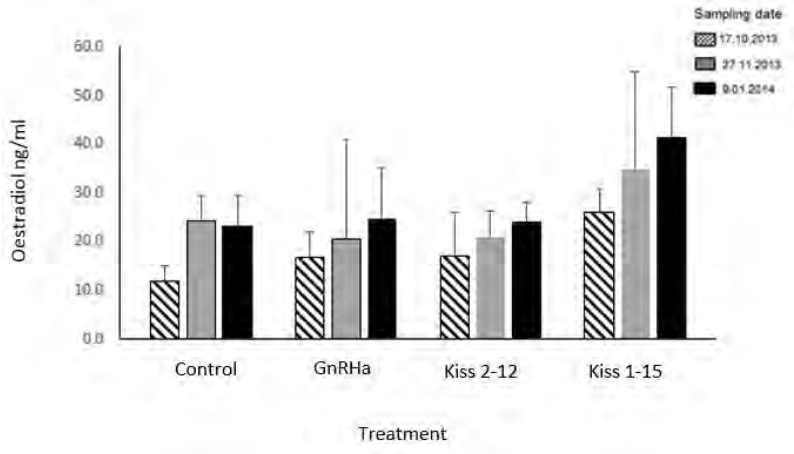


FIGURE 29. Mean plasma oestradiol (E_2) levels of female YTK implanted with blank implants (Control), GnRHa (50 µg/kg), and kisspeptin peptides (Kiss1-15; Kiss2-12; 50 µg/kg).

6.2.2 Recombinant YTK FSH (recYTK_FSH)

Another approach we trialled in order to shorten generation time was the use of recombinant YTK FSH. There are studies showing that recombinant FSH can stimulate maturation in prepubertal male European seabass (Mazón et al. 2013), goldfish (Kobayashi et al. 2006) and Japanese eel (Kazeto et al. 2014).

We generated recYTK_FSH according to Aizen et al. (2007). Full coding sequences of YTK FSH β and α subunits were first isolated by molecular cloning. A 6His-tagged single chain construct (recYTK_FSH) was synthesised and the recombinant peptide expressed in *P pastoris*. Western blot analysis of deglycosylated recombinant peptide showed the expected 23 kDa band (Figure 30). A high expressing clone was identified based on antibiotic resistance (Figure 31).

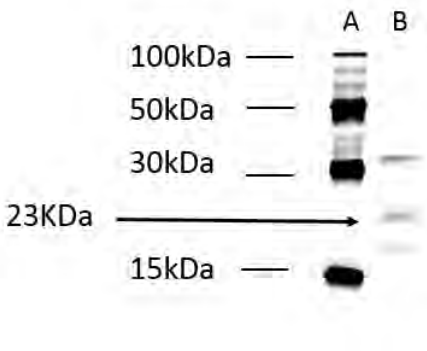


FIGURE 30. Western blot analysis of deglycosylated single chain recYTK_FSH showing the expected 23 kDa band (lane B) using antibodies to the 6His tag. (Lane A - protein ladder)

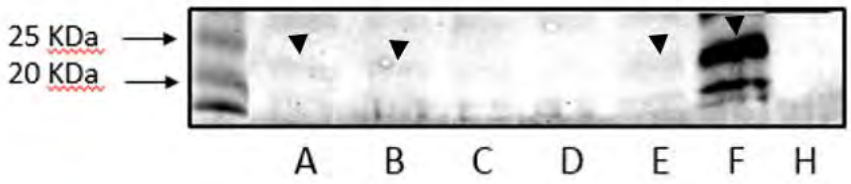


FIGURE 31. Western blot analysis of the recYTK_FSH obtained from 6 clones. The expected 23kDa band (arrowhead) was observed from clones A, B, E and F. Clone F clearly showed the highest yield of the recombinant protein. H is negative control (blank yeast).

Biological activity of the recYTK_FSH was confirmed using *in vitro* and *in vivo* bioassays. Using a sensitive reporter gene assay (CRE-luciferase as reporter) and a mammalian cell line, activation by the recombinant peptide of the tilapia FSH receptor was observed (Figure 32). *In vitro* bioassay using ovarian fragments from tilapia showed a dose response trend, in terms of E_2 secretion, however the differences were not significant (Figure 33). In contrast, E_2 release from YTK ovarian tissues showed a significant dose response (Figure 34). Similarly, 11KT secretion from YTK testicular tissues treated with increasing doses of recYTK_FSH also showed a significant positive response (Figure 34).

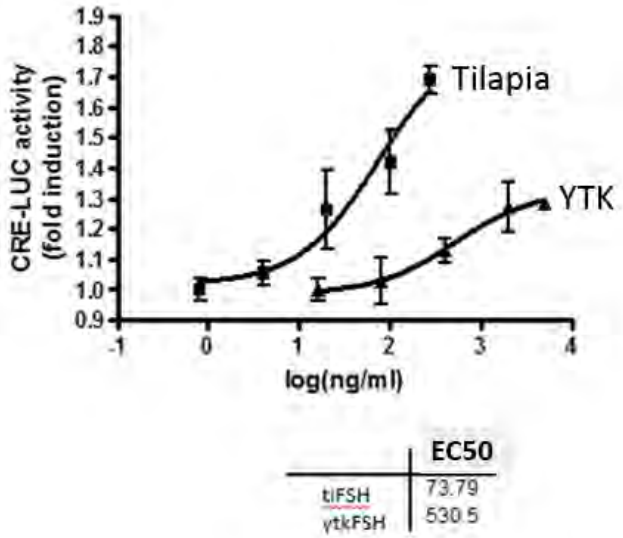


FIGURE 32. *In vitro* bioassay of recYTK_FSH (ytkFSH) and recombinant tilapia FSH (tiFSH) binding to the tilapia FSH receptor.

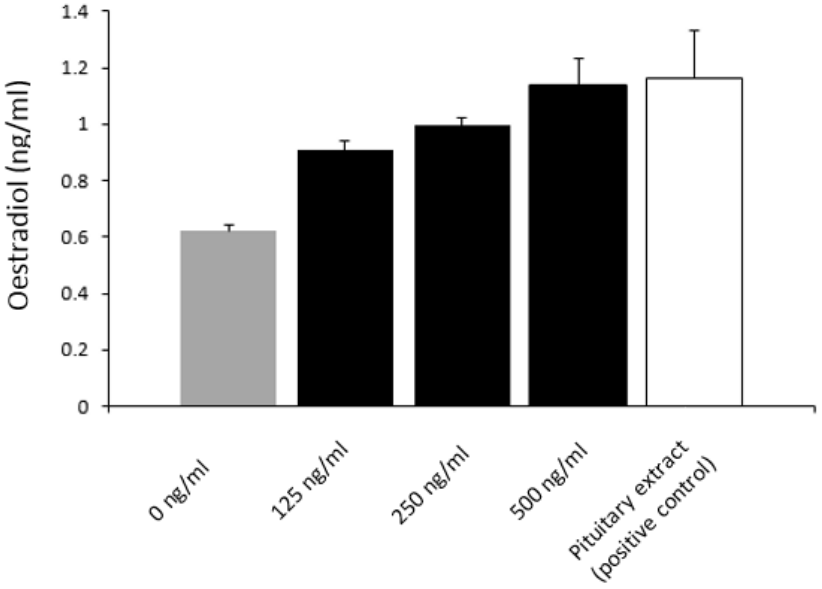


FIGURE 33. Secretion of E_2 (oestradiol) from tilapia ovarian fragments in response to various doses (X-axis) of recYTK_FSH. Diluted tilapia pituitary extract were used as a positive control. Culture media containing no hormone (0 ng/ml) was used as a negative control.

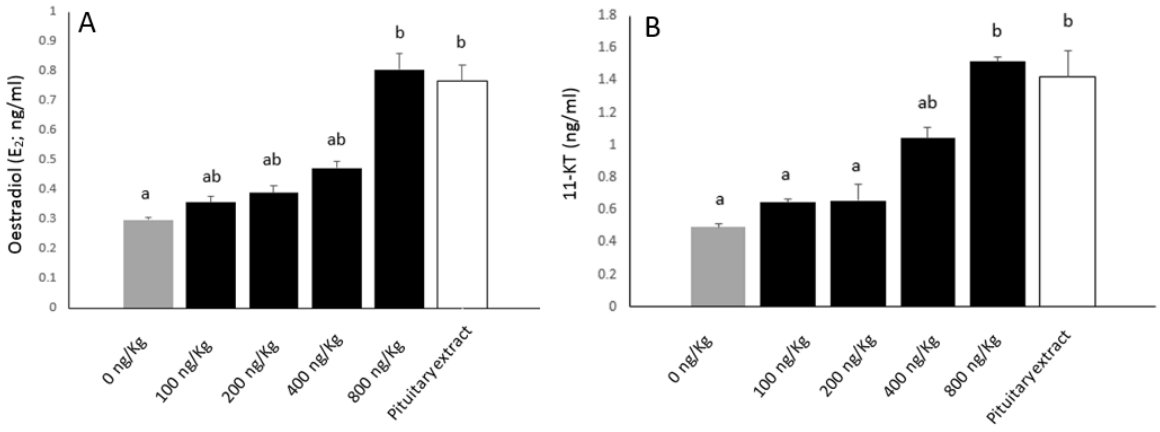


FIGURE 34. E_2 secretion from YTK ovarian fragments (A) and 11KT release from YTK testicular fragments (B) incubated *in vitro* with increasing doses (X-axis) of recYTK_FSH. Different letters indicate significant difference ($p<0.05$). The positive control (pituitary extract) was pituitary homogenate only.

In vivo assays in sexually immature female YTK resulted in a significant increase in plasma E_2 in treated fish compared with the control (Figure 35a). Mean oocyte diameter was higher in the treated group, but the difference with the control was not significant (Figure 35b). In males, plasma levels of 11KT was significantly lower in the group injected with recYTK_FSH compared with the control (Figure 36a), however a higher number of fish were spermiating in the treated group (Figure 36b).

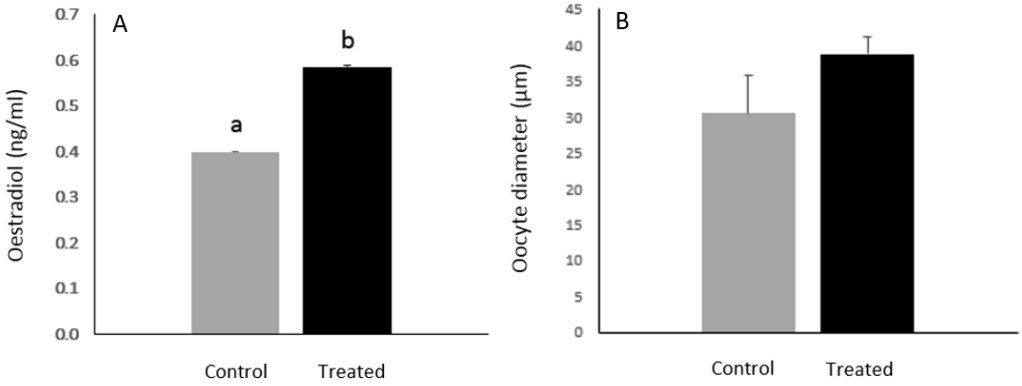


FIGURE 35. Mean plasma E_2 levels (A) and mean oocyte diameter (B) in female YTK treated *in vivo* with recYTK_FSH. Different letters indicate significant difference ($p<0.05$).

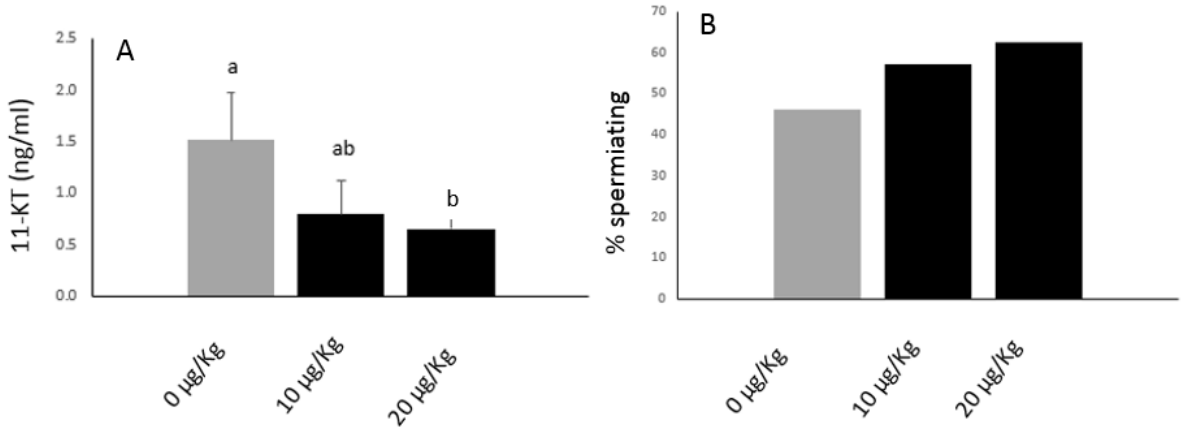


FIGURE 36. Mean plasma 11KT (A) and percentage spermiating fish (B) in male YTK treated *in vivo* with recYTK_FSH (doses shown on the X-axis). Different letters indicate significant difference ($p<0.05$).

Histological analysis of the cross section from the middle area of the right and left gonads showed no differences between the two regions, either in male or in female fish. When comparing the treated and untreated fish, recYTK_FSH-treated females (**Figure 37B**) appeared to exhibit more developed gonads than control group (**Figure 37A**). However, only previtellogenic oocytes were detected in both of the groups, with both still being in stage 1 of gonadal development.

In males, gonads from treated fish (**Figure 37D**) were more developed than gonads from the control group (**Figure 37C**). Spermatogonia, spermatocytes and spermatids were present in all of the YTK males. However, spermatozoa was only present in some of the animals belonging to the treated group (stage 2-3; **Figure 37D**).

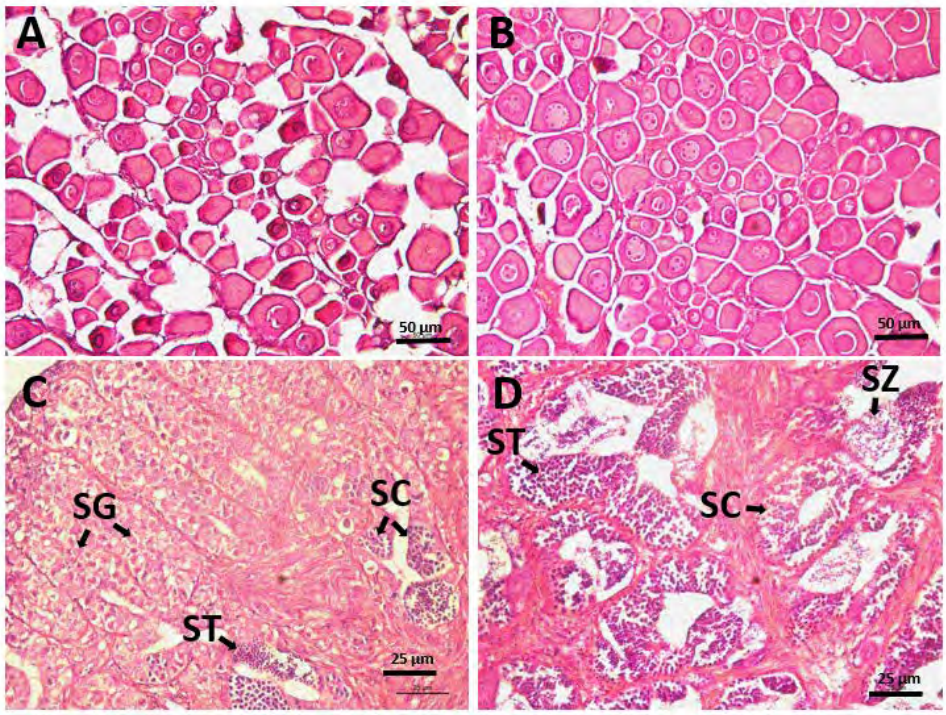


FIGURE 37. Histological sections of ovaries (A, B) and testis (C, D) of YTK treated with saline only (A, C) and with recYTK_FSH (B, D). Gonads were at stage 1 (A, B), stage 2 (C) and stage 3 (D) of reproductive maturation. (S= spermatogonia; SC= spermatocyte; ST= spermatids; SZ= spermatozoa) Stages 1 and 2 are classified as immature and stages 3 to 5 as mature (Poortenaar et al. 2001).

The *in vitro* and *in vivo* bioassays overall proves that our recombinant YTK FSH is biologically active. In YTK *in vivo*, the higher levels of E_2 in the plasma in the treated females is a good indication that our recombinant protein promoted reproductive development at the doses (10-20µg/kg) tested. Elevating the E_2 plasma levels is a pre-requisite to advancing reproductive development and this is a very encouraging result. When looking at the oocyte size in the control and treated females, although statistically significant differences were not found, there was a trend showing the oocytes from recYTK_FSH treated fish were bigger. It was unfortunate that only 5 treated females were in the cohort of fish we trialed (as we could not sex the fish at the start of the trial).

In contrast to the results obtained in female YTK, in males, lower levels of 11KT were observed in fish treated with recombinant FSH when compared with the control animals, however the level of spermiation in the treated fish was higher, showing a clear dose response. These results are consistent with 11KT levels dropping once spermiation commences.

We were quite limited with the number of fish for the *in vivo* trial, the duration of treatment, as well the tank facilities. However, our current results prove that our recYTK_FSH is indeed biologically active. Therefore, there is an opportunity to utilize the recombinant FSH for broodstock management. As the availability of recYTK_FSH would

facilitate developing tools to precisely define the timing of the activation of the brain-pituitary axis during sexual maturation, the administration of reproductive interventions could also be targeted at the most appropriate time point. We are setting up a biofermentor system for the production of recombinant peptides on a larger scale.

6.2.3 Oral delivery

We have been developing a system to deliver hormones, peptides and DNA constructs via the oral route. This is aimed towards minimising the handling stress in broodstock associated with injections or implantations that are currently used. We are exploring the use of chitosan-based encapsulation methods to protect the material to be introduced from degradation while in the digestive tract. The principles are similar to those being utilised in the delivery of therapeutic proteins manufactured by biotechnology (reviewed by Fonte et al., 2012). There is great interest in chitosan due to its several exceptional and unique properties. It is biodegradable, biocompatible, mucoadhesive, and presents low toxicity and low antigenic potential. In addition, it has been reported to have antimicrobial, antioxidant, analgesic and antitumoral activity (reviewed by Fonte et al., 2012).

In the present study, encapsulation of GnRHa was achieved either as liposomes or chitosan nanoparticles (100 nm diameter). ELISA analysis showed that incorporation of the GnRHa peptide can range from 60-70%. We are currently testing whether chitosan GnRHa nanoparticles that have been mixed in a formulated diet would be able to stimulate spawning in mature tilapia. The trial is at the very early stages, however we have observed spawning (3 out of 3 females) in the group fed with the encapsulated GnRHa.

We have generated two recombinant SBT luteinizing hormone (recSBT_LH) expression constructs, one linked with glycine and the other linked by carboxy terminal peptide (CTP) of the human chorionic gonadotropin. The constructs were engineered into the pcDNA3.1 expression vector, which is suitable for expression *in vivo*. Oral delivery of the CTP-linked construct, encapsulated in liposome, is being currently tested at the laboratory of Prof Ming-Wei Lu at the National Taiwan Ocean University. If proven effective, this may provide another broodstock management technique particularly for stimulating spawning, which is known in fish to be mainly regulated by LH (Yaron et al. 2003). There is evidence that DNA expression constructs incorporated in the diet can be absorbed from the gut and expressed *in vivo* (Ramos et al. 2005). In the European sea bass, *in vivo* expression of LH DNA expression construct, administered by injection, resulted in more prolonged circulating level of the hormone (Mazón et al. 2013). The technique we are exploring avoids the stress associated with multiple handling in case chronic treatment is required.

In addition to encapsulated compounds for oral delivery, we have also generated a recombinant peptide (recSBT_GnRH1) produced intracellularly in yeast for delivery through the diet. There is evidence that recombinant peptides expressed and contained in yeast, which is then delivered through the digestive system, can enter the peripheral circulation. This has been demonstrated in mice where yeast that expressed a recombinant antigen was mixed with the diet and successfully immunised the fed mice (Kim et al. 2014). We have already verified expression of the mature GnRH decapeptide by MALDI TOF/TOF mass spectrometry, which revealed a peak corresponding to the estimated 1.1 KDa molecular

weight of the GnRH1 decapeptide (Figure 38). We are planning directly sequencing the putative GnRH1 decapeptide to confirm its amino acid sequence and conduct *in vivo* experiments to verify biological activity.

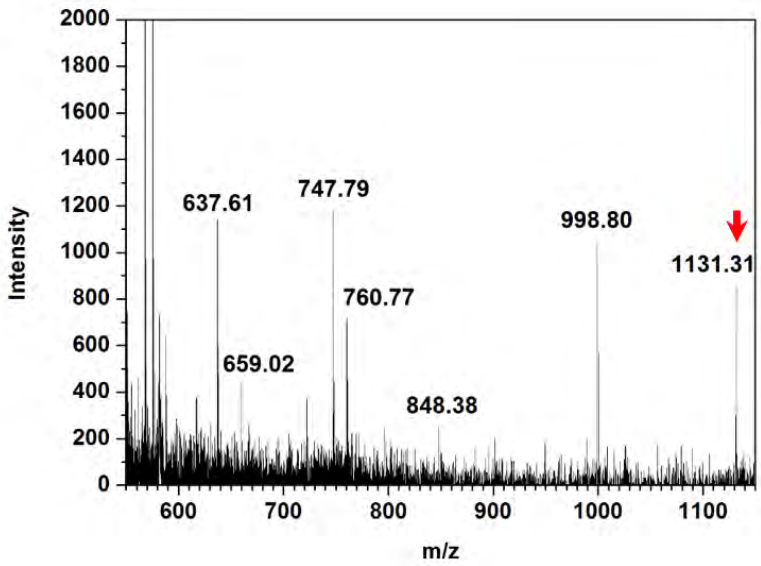


FIGURE 38. MALDI TOF/TOF mass spectrometry of total protein extract from yeast (*P pastoris*) transformed with recombinant SBT GnRH1. The red arrow indicates the putative recombinant SBT GnRH1 with a molecular weight of 1.1 kDa.

6.2.4 Transcriptomics: YTK Liver

A total of 17,200,694 raw paired-end reads from YTK liver were generated using the Illumina MiSeq sequencing platform. A slightly higher proportion of these reads (58.5 %) was from male fish with the remaining reads (41.4 %) from the female contribution. After quality screening, the surviving 11,998,721 high-quality PE reads were used for *de-novo* assembly.

Output from the assembler was screened for redundant and low complexity sequences producing a transcriptome of 199,751 non-redundant transcripts with an average length of 388.08 nucleotides (nt), an N50 of 733 nt and a total length of 77.51 megabases (Mb) (Table 8). Peptide prediction using the non-redundant nucleotide sequences from the *de-novo* assembly using Transdecoder software produced a set of 91,001 protein sequences with a maximum size of 4,283 amino acids and a mean length of around 102 amino acids.

In order to provide an overview of the gene content with respect to functional classifications within the transcriptome assembly, protein sequences were aligned against the NCBI eukaryotic orthologous groups of proteins (KOG) database (Figure 39). A total of 21,962 predicted protein sequences return a hit to this database. The most highly represented was the T category (signal transduction mechanisms) with 5,207 sequences reporting a match to this functional class.

Table 8. Statistics of *de-novo* assembly of the YTK liver transcriptome and overview of peptide prediction and annotation results.

Sequencer output	
Total raw PE reads male	10,066,194
Total raw PE reads female	7,134,500
Total clean PE reads for assembly	11,998,721 (69.75%)
Assembly	
Total non-redundant transcripts	199,751
Total consensus length (nt)	77,518,413
Largest transcript	14,281
Mean transcript length (nt)	388.08
N50 of transcripts	733
% GC content	48
Peptide Prediction and Annotation	
Number predicted proteins	91,001
Longest ORF	4,283
Mean ORF length	102.81
Number of sequences with KOG hits	21,962
Number of sequences with Pfam matches	29,151

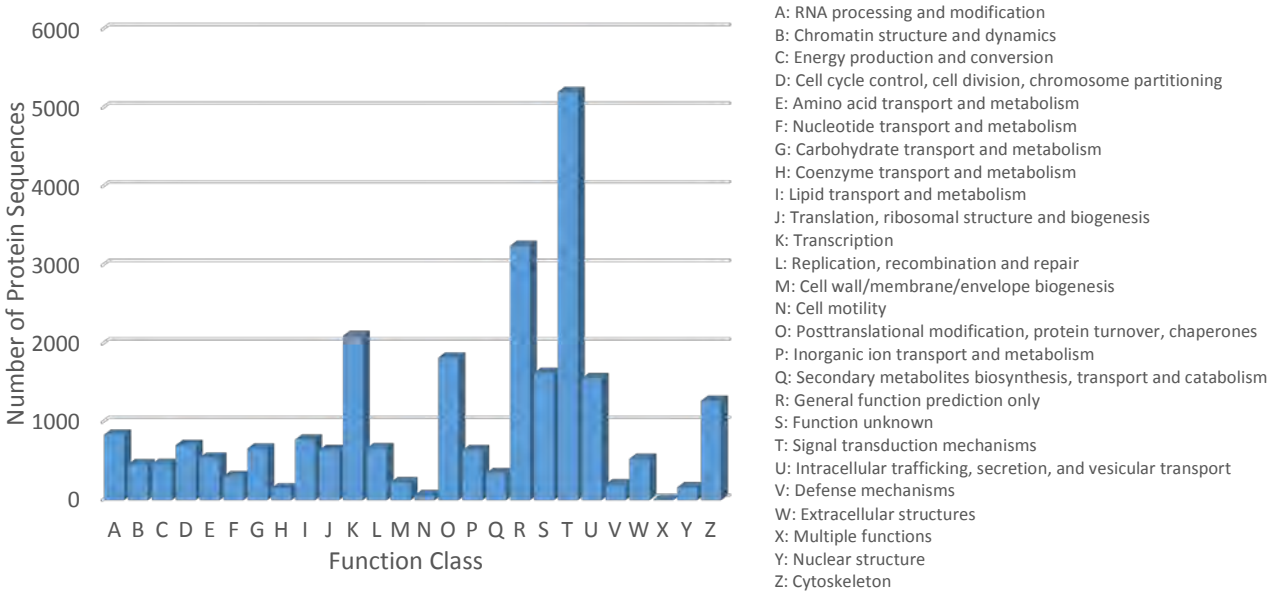


FIGURE 39. Distribution of predicted protein sequences annotated from the YTK liver transcriptome using the KOG database.

The predicted protein sequences were searched for matches to protein family domains of the Pfam database by alignment to hidden Markov models. A total of 29,151 protein sequences returned a match to approximately 4,800 unique Pfam entries. Predicted protein sequences were searched for genes of interest to reproduction within the Pfam and KOG analysis descriptors. Sequences were identified resembling full-length matches to isoforms of vitellogenin and oestrogen receptors alpha and beta.

We are in the process of generating transcriptome libraries from gonad and pituitary of mature female YTK. The transcriptome sequences can provide a snap shot of genes that are relevant for reproduction as well potential markers for specific traits.

6.2.5 References

Aizen J, Kasuto H, Golan M, et al. 2007. Tilapia follicle-stimulating hormone (FSH): immunochemistry, stimulation by gonadotropin-releasing hormone, and effect of biologically active recombinant FSH on steroid secretion. Biol Reprod 76: 692-700.

Biran J, Ben-Dor S, Levavi-Sivan B. 2008. Molecular identification and functional characterization of the kisspeptin/kisspeptin receptor system in lower vertebrates. Biol Reprod 79: 776-786.

Bolger A, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 1–7.

Finn RD et al. 2014. Pfam: the protein families database. Nucleic Acids Res 42:D222–30.

Fonte P, Andrade JC, Seabra V, Sarmento B. 2012. Chitosan-based nanoparticles as delivery systems of therapeutic proteins. In: Voynov V and Caravella JA (eds.), Therapeutic Proteins: Methods and Protocols, Methods in Molecular Biology, vol. 899, DOI 10.1007/978-1-61779-921-1_28, © Springer Science+Business Media, LLC.

Haas BJ et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494–512.

Invitrogen. 2010. Pichia Expression Kit User Manual.

Kafka AP, Rades T, McDowell A. 2010. Rapid and specific high-performance liquid chromatography for the *in vitro* quantification of D-Lys^6-GnRH in a microemulsion-type formulation in the presence of peptide oxidation products. Biomed Chromatog 24: 132-139.

Kazeto Y, Ito R, Ozaki Y, et al. 2014. Mass production of Japanese eel recombinant follicle-stimulating hormone and luteinizing hormone by a stable expression system: they fully induced ovarian development at a differential mode. Presented at the 10th International Symposium on the Reproductive Physiology of Fish, 25-30 May 2014, Olháo, Portugal.

Kobayashi M, Morita T, Ikeguchi K, et al. 2006. In vivo biological activity of recombinant

goldfish gonadotropins produced by baculovirus in silkworm larvae. Aquaculture 256: 433-442.

Kim H J, Lee JY, Kang HA, Lee Y, Park EJ, Kim HJ. 2014. Oral immunization with whole yeast producing viral capsid antigen provokes a stronger humoral immune response than purified viral capsid antigen." Lett Applied Microbiol 58: 285-291.

Mazón MA, Gómez A, Yilmaz O et al. 2014. Administration of follicle-stimulating hormone in vivo triggers testicular recrudescence of juvenile European sea bass (*Dicentrarchus labrax)*. Biol Reprod 90: 1-10.

Mazón MJ, Zanuy S, Muñoz I, et al. 2013. Luteinizing hormone plasmid therapy results in long-lasting high circulating Lh and increased sperm production in European sea bass (*Dicentrarchus labrax*). Biol Reprod 88: 1-7.

Molés G, Zanuy S, Muňoz I, et al. 2011. Receptor specificity and functional comparison of recombinant sea bass (*Dicentrarchus labrax*) gonadotropins (Fsh and Lh) produced in different host systems. Biol Reprod 84: 1171–1181.

Mylonas, C.C., Bridges, C.R., Gordin, H., et al. 2007. Preparation and administration of gonadotropin-releasing hormone agonist (GnRHa) implants for the artificial control of reproductive maturation in captive-reared Atlantic bluefin tuna (*Thunnus thynnus thynnus*). Rev Fish Sci 15, 183–210.

Nocillado JN, Zohar Y, Biran J, Levavi-Sivan B, Elizur A. 2013. Chronic kisspeptin administration stimulated gonadal development in pre-pubertal male yellowtail kingfish (*Seriola lalandi*: Perciformes) during the breeding and non-breeding season. Gen Comp Endocrinol 191: 168-176.

Nocillado JN, Zohar Y, Biran J, Levavi-Sivan B, Mechaly S, Zohar Y, Elizur A. 2012. The Kiss2 receptor (Kiss2r) gene in Southern Bluefin Tuna, *Thunnus maccoyii* and in Yellowtail Kingfish, *Seriola lalandi* – functional analysis and isolation of transcript variants. Mol Cell Endocrinol 362: 211-220.

Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8: 785-786.

Ramos EA, Relucio JLV, Torres-Villanueva CAT. 2005. Gene expression in tilapia following oral delivery of chitosan-encapsulated plasmid DNA incorporated into fish feeds. Mar Biotechnol 7: 89-94.

Wu S, Zhu Z, Fu L, Niu B, Li W. 2011. WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genomics 12:444.

Yaron Z, Gur G, Melamed P, et al. 2003. Regulation of fish gonadotropins. Int Rev Cytol 225: 131-185.

7 Benefits and Adoption

Today, all CST YTK production is based on the offspring of selected parents, i.e. **they are all F_2 fish**. All evidence and anecdotal information is that the selected stock is superior to the wild stock. We can say then the work of this project relating to YTK genetics was completely adopted, and the benefits look like being in the order of 20% or more per generation.

Also, the project has delivered a swath of high quality DNA microsatellite markers that are fully adopted, and beneficial for, the operation of the ongoing selection program.

Overall a YTK breeding plan was implemented with expected genetic gains in culture performance of fish produced in the 2013-4 season of $1,000,000 /1000 ton (i.e. @10% gain). If production is just 1,000 tons, this represents a 8:1 return for just one year of production on investment of $120,000 in the genetics over the four years of the project (minus extra feed costs). If the production continues for say 10 years, or if production is larger, or if genetic gains increase each generation, the return on this particular investment could be ten times more than 8:1, minus any new investment made in the genetics. Even if the cost benefit values are upper levels, and the cost of money is not considered, the overall calculations are quite robust, suggesting here, as always, whenever organizations/companies commit to formal genetic improvement and sustains the program to maturity, the investment is returned many many times over.

8 Conclusion

CST is now enjoying the production benefits from having invested and sustained over the years a genetic improvement program for YTK. This program, sustainable both in terms of cost and Ne, will not only continue to support the profitability of CST, but may also open new markets in the export of genetically improved fingerlings worldwide.