Development of a porcine skeletal muscle cDNA microarray: analysis of differential transcript expression in phenotypically distinct muscles
Qianfan Bai1, Christine McGillivray1, Nuno da Costa1, Saffron Dornan2, Gary Evans2, Michael James Stear1 and Kin-Chow Chang*1

Address: 1Laboratory of Veterinary Molecular Medicine, Department of Veterinary Pathology, University of Glasgow, Bearsden Road, Glasgow G61 1QH, UK and 2Sygen International, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK
Email: Qianfan Bai - 9912270b@student.gla.ac.uk; Christine McGillivray - cph1u@udcf.gla.ac.uk; Nuno da Costa - nadc1a@udcf.gla.ac.uk; Saffron Dornan - scd@mole.bio.cam.ac.uk; Gary Evans - evans@fyf.pig.co.uk; Michael James Stear - m.j.stear@vet.gla.ac.uk; Kin-Chow Chang* - k.chang@vet.gla.ac.uk
* Corresponding author

Published: 1 March 2003
Received: 5 December 2002
Accepted: 1 March 2003

Abstract

Background: Microarray profiling has the potential to illuminate the molecular processes that govern the phenotypic characteristics of porcine skeletal muscles, such as hypertrophy or atrophy, and the expression of specific fibre types. This information is not only important for understanding basic muscle biology but also provides underpinning knowledge for enhancing the efficiency of livestock production.

Results: We report on the de novo development of a composite skeletal muscle cDNA microarray, comprising 5500 clones from two developmentally distinct cDNA libraries (longissimus dorsi of a 50-day porcine foetus and the gastrocnemius of a 3-day-old pig). Clones selected for the microarray assembly were of low to moderate abundance, as indicated by colony hybridisation. We profiled the differential expression of genes between the psoas (red muscle) and the longissimus dorsi (white muscle), by co-hybridisation of Cy3 and Cy5 labelled cDNA derived from these two muscles. Results from seven microarray slides (replicates) correctly identified genes that were expected to be differentially expressed, as well as a number of novel candidate regulatory genes. Quantitative real-time RT-PCR on selected genes was used to confirm the results from the microarray.

Conclusion: We have developed a porcine skeletal muscle cDNA microarray and have identified a number of candidate genes that could be involved in muscle phenotype determination, including several members of the casein kinase 2 signalling pathway.

Background
Skeletal muscle is a heterogeneous tissue that has the ability to rapidly undergo biochemical and physical changes in response to external stimuli, such as appropriate nervous and hormonal stimulations, to adapt to the accompanying functional demands imposed on it. There is wide variation in phenotype between anatomical muscles in the body. Postural muscles, often described as red muscles, such as the soleus and psoas in the pig, are under continuous but modest activation. They are adapted to
undertake chronic contractile activity without fatigue, under aerobic respiratory conditions. On the other hand, white muscles, such as the *gastrocnemius* and *longissimus dorsi*, are recruited sporadically during brief periods of intense muscular activity, like running. They are susceptible to fatigue as their main source of energy is derived from anaerobic glycolysis. Red muscles are better endowed with capillaries, myoglobin, lipids and mitochondria than white muscles [1,2]. Red and white muscles also differ in their fibre type composition. Postnatal mammalian muscles (e.g. in pigs and rodents) have 4 major fibre types characterised by the expression of the slow/β, 2a, 2x and 2b myosin heavy chain (MyHC) gene isoforms [2–4]. Based on MyHC isoform expression, postnatal muscle fibres in the pig have recently been resolved by the combination of immunocytochemistry and *in situ* hybridisation into 4 major types [5–7]. Metabolic, biochemical and biophysical characteristics, such as oxidative and glycolytic capacities, fibre size, colour, and glycogen and lipid contents, vary between MyHC fibre types [2,8,9]. The slow/β and fast 2b fibres, also known as slow oxidative and fast glycolytic respectively, represent two extreme metabolic profiles. The fast 2a and fast 2x fibres are intermediate fast oxidative-glycolytic fibres. Fast 2a fibres are more closely related to slow/f fibres, and fast 2x are more similar to fast 2b fibres. Hence the composition of fibre types in a muscle is a major determinant of its phenotypic properties.

Understanding the molecular processes that govern the phenotypic characteristics of muscles, such as hypertrophy/atrophy, and expression specific fibre types, is of agricultural and medical importance [9–12]. Microarray technology can simultaneously measure the differential expression of a large number of genes in a given tissue and may identify the genes responsible for the relevant phenotype [13]. We report here on the *de novo* development of a porcine skeletal muscle cDNA microarray, comprising 5500 clones from two cDNA muscle libraries. Its functional integrity was assessed by profiling the differences in gene expression between red (*psoas*) and white (*longissimus dorsi, LD*) muscles. Among the differentially expressed genes, a number of novel candidate genes were found that could determine the phenotypic differences among different muscle types.

### Results

**Construction of a composite porcine skeletal muscle cDNA microarray**

The *de novo* development of a porcine skeletal muscle cDNA microarray was based on the use of two representative, directionally cloned *ZAP-Express* cDNA expression libraries; one derived from the back muscle of a 50-day-old foetus and the other from the *gastrocnemius* of a 3-day-old pig. Genes that are expressed in skeletal muscle are likely to be present in one or both libraries. In our porcine skeletal muscle cDNA libraries, about a third of the clones were found to be from highly expressed genes, based on signal intensity of hybridisation (Fig. 1A). Weakly to moderately expressed genes were selected for microarray assembly; highly expressed clones were avoided (Fig. 1B). On average, between a quarter and a half of the randomly picked colonies from each agar plate were eventually selected for microarray assembly. A total of 5,500 plasmid clones were selected, of which 3,500 clones were taken from the 3-day-old muscle library and 2,000 clones were taken from the 50-day foetal muscle library. After each production step of miniprep plasmid preparation, insert amplification, and PCR product purification, about a quarter of the selected clones were checked by gel electrophoresis or spectrophotometry. Approximately 10% of all selected clones were found to be without any cDNA insert, a figure expected in a typical cDNA library (data not shown). Each selected clone was printed twice on the microarray (Fig. 2A).

**Functional assessment of the porcine cDNA microarray: red-white muscle analysis**

To evaluate the performance of our composite porcine skeletal muscle cDNA microarray, a profiling experiment was conducted to determine the differential expression of genes between red (*psoas*) and white (*longissimus dorsi, LD*) muscle. Dual-colour hybridisation (Fig. 2A) was performed on seven replicate microarray slides. An intensity-dependent (LOWESS) step was used to normalise data. Fig. 2B plots the median expression of each clone in the *psoas* against the median expression of the same clone in the LD muscle. Each point is the median of 14 values (2 replicates per slide, 7 slides per muscle). Most points cluster around the middle line indicating similar levels of expression in both muscles. There were, however, a number of clones falling substantially below the line, indicating consistently lower levels of expression in the *psoas* compared with the LD muscle. Low signal readings for both dyes may indicate the absence of a cDNA insert. About 12% of the printed clones were found to fall below 50 units for both dyes and were considered to be without cDNA inserts. This value was consistent with the earlier estimate of insertless clones in the cDNA libraries.

A normalised *psoas*/LD ratio of 2.0 or more was used to identify genes that were more highly expressed in *psoas* than in LD muscle. This ratio represents the top 5% of differences in expression. Seventy clones meeting this criterion were sequenced. The sequences were compared with database sequences by BLAST searching. Table 1 is a summary of these genes. Sixty seven percent of the clones (47 out of 70) sequenced were genes of mitochondrial origin. Of these several were featured more than once on the microarray, namely genes encoding 16S ribosomal RNA, 12S ribosomal RNA, and NADH dehydrogenase subunits.
3 and 6. Thirty percent of the clones (21 out of 70) did not show any homology with known mitochondrial or sarcormeric genes. The function of 9 of these clones was completely unknown. One gene encoded for fructose-1,6-biphosphatase, an enzyme that is necessary for muscle gluconeogenesis. The function of the remaining 11 clones were involved with aspects of transcription, translation or signal transduction, but their functions in skeletal muscle have not been characterised. Several genes were members of the casein kinase 2 signalling pathway. The α1 subunit of casein kinase 2 (CK2) is one half of the holoenzyme [14,15]. The small muscle protein (smpx) is encoded by a recently discovered X-linked stretch response gene [16].

The tyrosine kinase A6-related protein binds ATP and actin, and interacts with protein kinase C zeta [17]. Although the function of the latter two proteins is not known, both were shown to be targets of CK2 phosphorylation.

Clones with a normalised psoas/LD ratio of 0.7 or less, which represented the most extreme 5% of clones were identified. Forty-five clones were sequenced and examined for homology by BLAST searching. Table 2 lists the genes that were more highly expressed in the LD than in the psoas. Only 4 out of 45 clones were of mitochondrial origin, and they were different from those expressed more
abundantly in the psoas (Table 1). Fast isoforms of sarcomeric proteins (myosin heavy chains [MyHCs] 2a, 2x, 2b, myosin regulatory light chain 2, α-actinin 3, fast troponin C, and fast troponin T3) were well represented amongst the sequenced clones. Sarcomeric/structural genes made up nearly half of the total clones (22 out of 45). The other highly represented group of genes on the list (11 out of 45 clones) were involved in glycolysis, such as glyceraldehyde 3-phosphate dehydrogenase (GAPDH). The function of two non-sarcomeric genes was completely unknown. A few candidate regulatory genes (bin1, polyubiquitin, myomegalin-like, and HSPA8) were also found. Of particular interest is the tumour suppressor gene bin1 [18], which has recently been shown to play a role in C2C12 myoblast differentiation [19].

**Quantitative expression of selected genes: validation by TaqMan real time RT-PCR**

To assess the validity of the microarray approach to identify differentially expressed genes, quantitative real-time RT-PCR (TaqMan) was performed on four representative clones, GAPDH, MyHC 2b, bin1, and a novel gene (kc2725), each normalised to β-actin. In line with functional expectations, recent quantitative work performed has shown that the mRNA expression of GAPDH and MyHC 2b was higher in the psoas than in the LD within individual animals [20]. In this study, the relative levels of GAPDH and MyHC 2b of pooled total cDNA samples from three 22-week-old Berkshire pigs were measured (Fig. 3). Our results showed that the differential expression of GAPDH and MyHC 2b was sufficiently consistent between the two muscles to be detected in pooled cDNAs. A novel gene kc2725 of unknown function, found by
Table 1: Genes more highly expressed in psoas than in LD

| Gene                                      | Psoas/LD | Homology | # of clones | Accession # |
|-------------------------------------------|----------|----------|-------------|-------------|
| **Mitochondrial origin**                  |          |          |             |             |
| 16S ribosomal RNA                          | 2.07     | 97% (S)  | 12          | AY011178    |
| 12S ribosomal RNA                          | 2.18     | 100% (S) | 7           | AY012145    |
| 28S ribosomal RNA                          | 2.67     | 99% (R)  | 1           | V01270      |
| Ribosomal protein S29                      | 4.14     | 94% (B)  | 1           | U66372      |
| Ribosomal protein L23                      | 30.05    | 92% (H)  | 1           | NM_000978   |
| Cytochrome oxidase subunit I              | 2.03     | 86% (B)  | 2           | AF490529    |
| Cytochrome oxidase subunit III             | 2.28     | 81% (T)  | 1           | AF030272    |
| Cytochrome oxidase subunit Vlc             | 27.97    | 88% (H)  | 1           | NM_004374   |
| Cytochrome oxidase subunit Vlb             | 2.51     | 96% (B)  | 1           | X15112      |
| NADH dehydrogenase subunit 2              | 2.51     | 77% (P)  | 1           | AF414121    |
| NADH dehydrogenase subunit 3              | 2.58     | 82% (B)  | 7           | AY052631    |
| NADH dehydrogenase subunit 6              | 2.15     | 81% (B)  | 4           | AF416451    |
| NADH dehydrogenase subunit B15            | 2.34     | 87% (B)  | 1           | X64898      |
| ATP synthase protein 6                     | 2.1      | 96% (S)  | 2           | AF190813    |
| ATP synthase protein 8                     | 2.17     | 100% (S) | 1           | AF039170    |
| ATP synthase protein 9                     | 20.55    | 96% (H)  | 1           | NM_001689   |
| tRNA                                       | 3.87     | 97% (S)  | 2           | AF304202    |
| Adenine nucleotide translocator SLC25A6    | 2.05     | 92% (H)  | 1           | XM_114724   |
| **Sarcomeric / structural**                |          |          |             |             |
| Ankyrin I                                  | 2.23     | 93% (H)  | 1           | XM_016774   |
| Tubulin alpha 1                            | 2.68     | 94% (H)  | 1           | AF141347    |
| **Non-sarcomeric**                         |          |          |             |             |
| Pituitary tumor-transforming 1 interacting protein (tr) | 2.98 | 84% (H) | 1 | BC031097 |
| Par-6 partitioning defective 6 homolog gamma (tr) | 30.36 | 92% (M) | 1 | XM_129044 |
| TAFII140, a novel TAF component (TATA-binding protein associated factors) (tr) | 2.16 | 94% (H) | 1 | AJ292190 |
| STAF transcriptional activator (tr)        | 2.15     | 88% (H)  | 1           | NM_003442   |
| Prefoldin subunit 6 (KE2 protein) (tl)     | 2.65     | 85% (M)  | 1           | M65255      |
| EH domain-binding protein, Epsin (s)       | 2.23     | 93% (H)  | 2           | NM_013333   |
| Beta-catenin (s)                           | 2.18     | 100% (S) | 1           | AB046171    |
| Casein kinase 2, alpha 1 polypeptide (s)   | 6.47     | 87% (H)  | 1           | NM_001895   |
| Small muscle protein (smpx) (s)            | 7.34     | 88% (H)  | 1           | BC005948    |
| Tyrosine kinase 9-like (A6-related protein) (s) | 2.04 | 90% (H) | 1 | BC016452 |
| Fructose-1,6-biphosphatase (m)             | 25.07    | 92% (O)  | 1           | AJ272520    |
| Tewis intracellular mediator protein (PEAS) (u) | 2.12 | 88% (H) | 1 | BC021546 |
| Hypothetical protein (FLU2666) (u)         | 2.23     | 93% (H)  | 1           | NM_024595   |
| Unknown small protein (u)                  | 3.93     | 94% (H)  | 1           | BC005398    |
| Unknown clone (kc2725), homologous to 3' untranslated region (UTR) of 16S rRNA binding protein S1 (u) | 2.57 | *90% (H) | 1 | *XM_028847 |
| Unknown clone, homologous to 3' UTR of enhancer of rudimentary (u) | 2.07 | *81% (H) | 1 | *U66871 |
| Unknown clone, homologous to Homo sapiens clone *KIAA0513 (u) | 13.21 | *81% (H) | 1 | *NM_014732 |
| Unknown clones (kc2668, kc522, kc2469) (u) | 3 | 3 | | |

Possible involvement: (tr), transcriptional; (tl), translational; (s), signalling; (m) etabolic; (u) unknown. Accession # refers to porcine gene or close homologue: (S) = sus scrofa, (R) = rattus norvegicus, (B) = bos taurus, (H) = homo sapiens, (T) = tragelaphus buxtoni, (P) = physeter catodon, (M) = mus musculus, (O) = ovis aries.
microarray analysis to be more highly expressed in the psoas than in the LD (Table 1), was individually quantified in 4 pigs (Fig. 4A). CK2 phosphorylation sites were also predicted in the deduced kc2725 protein (data not shown). All 4 pigs gave the same pattern of expression as that detected on the microarray. Interestingly, expression of kc2725 was ubiquitous and was much more abundant in other tissues than in skeletal muscle (Fig. 4B). However, the expression of a gene from the same muscle can vary between similar individuals (Fig. 4A and Fig. 5A).

Table 2: Genes more highly expressed in LD than in psoas

| Gene                                              | LD/Psoas | Homology | # of clones | Accession # |
|---------------------------------------------------|----------|----------|-------------|-------------|
| **Mitochondrial origin**                          |          |          |             |             |
| Ribosomal protein L35A                            | 2.2      | 86% (H)  | 1           | AK055653    |
| Ribosomal protein L7a                             | 1.85     | 90% (H)  | 1           | BC005128    |
| Ribosomal protein S23                             | 1.91     | 88% (R)  | 1           | X77398      |
| Ribosomal protein S12                             | 1.61     | 100% (S) | 1           | X79417      |
| **Sarcomeric / structural**                       |          |          |             |             |
| Myosin heavy chain fast 2a                        | 1.91     | 100% (S) | 6           | AB025260    |
| Myosin heavy chain fast 2x                        | 1.77     | 100% (S) | 2           | AB025262    |
| Myosin heavy chain fast 2b                        | 1.88     | 97% (S)  | 2           | AB025261    |
| Myosin regulatory light chain 2                    | 2.95     | 91% (H)  | 2           | AF363061    |
| Alpha actinin 3                                   | 1.47     | 93% (H)  | 2           | M86407      |
| S-nexilin                                         | 1.77     | 70% (H)  | 1           | AK057954    |
| Fast skeletal troponin C                          | 1.67     | 94% (O)  | 1           | Y00760      |
| Fast skeletal troponin T (TrT3)                   | 1.7      | 93% (B)  | 3           | AB085599    |
| Alpha-1 skeletal actin (ACTA1)                    | 1.53     | 93% (H)  | 2           | BC012597    |
| Alpha-2 smooth muscle actin (ACTA2)               | 1.62     | 96% (H)  | 1           | BC017554    |
| **Non-sarcomeric**                                |          |          |             |             |
| Sarcomplasmic/endoplasmic reticulum calcium ATPase 1 (SERCA1) (g) | 1.54 | 98% (S) | 3 | AY027797 |
| Aldolase A (g)                                    | 1.47     | 87% (H)  | 1           | BC010660    |
| Phosphoglycerate kinase (g)                       | 1.48     | 88% (H)  | 1           | V00572      |
| Glyceraldehyde-3-phosphate dehydrogenase (g)      | 1.53     | 100% (S) | 1           | X94251      |
| Muscle pyruvate kinase (PKM2) (g)                 | 1.63     | 91% (H)  | 1           | BC007952    |
| Phosphoglucomutase isomerase 2 (PGM) (g)          | 1.81     | 94% (O)  | 1           | M97664      |
| Muscle glycogen phosphorylase (g)                 | 1.65     | 92% (O)  | 2           | X04265      |
| Adenylate kinase isozyme 1 (myokinase) (g)        | 1.51     | 87% (S)  | 1           | E03007      |
| Calseenin (c)                                     | 1.7      | 83% (C)  | 2           | J03765      |
| Polyubiquitin (tl)                                | 1.51     | 90% (A)  | 1           | AF038129    |
| Bridging-integrator protein-1 (bin1) (s)          | 2.17     | 85% (H)  | 1           | AF001383    |
| Myomegalin-like protein (u)                       | 6.94     | 94% (H)  | 1           | AB042557    |
| Heat shock cognate protein (HSPA8) (u)            | 1.77     | 93% (B)  | 1           | X53827      |
| Unknown clone, homologous to Homo sapiens clone HSPC040 (u) | 2.08 | 86% (H) | 1 | BC000810 |
| Unknown clone, homologous to 3’ UTR of Homo sapiens α-actin bundling protein (u) | 1.99 | *77% (H) | 1 | *U09873 |

Associated or possible role: (g), glycolysis; (c), contraction; (tl), translation; (s), signalling; (u) unknown. Accession # refers to porcine gene or close homologue. Genus abbreviations as in Table 1.

Discussion

Development of porcine skeletal muscle cDNA microarray

We have constructed a composite porcine skeletal muscle cDNA microarray consisting of 5,500 clones from two de-
velopmentally distinct libraries, one derived from a 50-
day foetal longissimus dorsi muscle, and the other derived
from a 3-day-old gastrocnemius muscle. The choice of two
developmentally distinct libraries was to increase the
range of temporally-regulated genes represented in the
microarray, to extend its suitability for use in different
microarray-based muscle experiments.

The selection emphasis for inclusion on the microarray
was on lowly or moderately expressed clones. This clone
selection reduced the representation of highly expressed
genes, such as those encoding for sarcomeric proteins, and
increased the likelihood of including rare transcripts, in-
cluding those of regulatory importance. An alternative
method of normalising clone selection, which has been
used by others, is by reassociation of single-stranded li-
brary plasmids at relatively low Cot to remove highly ex-
pressed clones [22,23]. Another commonly used and
commercially available method of selection normalisa-
tion is subtractive suppression hybridisation [24]. We
chose not to use this method for microarray clone selec-
tion because of the possibility of excluding lowly ex-
pressed non-muscle specific clones.

A possible disadvantage of our porcine microarray, at least
at the beginning, is the lack of knowledge of the identity
of each clone. However, there is, at present, insufficient se-
quence information on farm animals to design a compre-
hensive oligonucleotide-based microarray. To date, we
have sequenced about 10% of our microarray clones (data
not shown). A major advantage of our microarray is that
we are likely to be in possession of the corresponding full
length cDNA clones, whose inserts were unidirectionally
cloned into a CMV-promoter driven expression plasmid
(pBK-CMV vector). These clones could be readily used for
downstream expression studies. As the identity of each
clone is unknown, we do not know for certain how repre-
sentative is the porcine skeletal muscle cDNA microarray.

On the one hand, considering that there might be fewer
than 30,000 human genes [25], and assuming that 50% of
all genes are transcriptionally active at one time in a given
tissue [26], it is possible that around 20% of the genes that
are expressed in skeletal muscle are found on our
microarray [25,27]. On the other hand, a group of 4080
human skeletal muscle genes, which included both skele-
tal muscle-specific genes and genes expressed in skeletal
muscle as well as in other tissues, was found to corre-
spond to 80% of the total number of genes expressed in
skeletal muscle as reported so far in Unigene (including
foetal muscle and rhabdomyosarcoma) [26]. Hence our
microarray may represent substantially more than 20% of
all genes expressed in porcine muscle.

Figure 3
TaqMan quantitative real-time RT-PCR analysis of GAPDH, and MyHC 2b mRNAs in porcine psoas (red) and LD (white) mus-
cles. Results of pooled total cDNA samples from three 22-week-old Berkshire pigs. GAPDH and MyHC 2b were about 2.7 and
2.5 times more highly expressed in the LD than psoas, respectively. Error bar = standard deviation.
Differential gene expression in red and white muscles

One objective of the red-white muscle experiment was simply to test the function of the newly assembled porcine microarray, from which two gene lists were generated (Tables 1 and 2). Genes that were expected to be differentially expressed and genes that were novel were found on each list. The microarray results validated our prior hypothesis of differential gene expression in red and white muscles, thus demonstrating the functional integrity of our newly constructed microarray. One of the well established distinguishing features of red muscle is its relatively high oxidative phosphorylation capacity, reflected by an abundance of mitochondria in red muscles. It is reassuring that genes from the mitochondrial genome were well represented in the red muscle pool of differentially expressed genes (Table 1). White muscles comprise predominantly more fast-glycolytic fibres than red muscles. Our findings were consistent with expectations, in that most of the 45 clones selected as more highly expressed in white muscle (Table 2) were either fast isoforms of structural genes, or enzymes connected with anaerobic glycolysis. With highly expressed genes, such as mitochondrial genes and structural genes, there were detectable levels of redundancies (repeats) on the microarray. However, in the presence of redundancies, no gene was found to be present on both lists. Differentially expressed clones originating from both the foetal and neonatal libraries were found in comparable proportion (data not shown). A recent report comparing red and white murine skeletal muscles on generic commercial oligonucleotide chips (Affymetrix

---

**Figure 4**

Tissue distribution of a novel gene kc2725. **A.** Expression of kc2725 was higher in the psoas than in the LD in all 4 pigs. Pigs 1 to 3 were 22-week-old Berkshires, and pig 4 was a 7-week-old cross breed. **B.** The expression of kc2725 was ubiquitous and much more abundant in other tissues. Error bar = standard deviation.
GeneChips), which comprised the equivalent of 3,000 different genes, yielded a differential list of 49 known genes [28]. Our results of known differentially expressed genes were comparable in both number and in the different types of genes found.

One problem underlying the analysis of microarray data is the large number of comparisons required, which can produce false positive results [29]. We compared the expression of clones in psoas and LD muscle as the median of 14 comparisons (each microarray clone was printed twice on each slide and 7 slides were used), and restricted further analysis to the clones demonstrating the most consistent and extreme differences. Therefore, the majority of our identified clones are likely to represent real differences between muscles. However, any specific gene assignment should be regarded as provisional until it can be confirmed in an independent study, such as another microarray or in another assay such as quantitative PCR.

The interpretation of microarray results can be complicated. Firstly, members of the same gene family could cross-hybridise to the same spots on the microarray. Interpretation of differential expression of individual isoforms should therefore be made with caution. In the case of MyHC genes (Table 2), a plausible interpretation is that fast MyHC mRNA isoforms (2a, 2x and 2b) were more abundant in the LD than in the psoas. However, we have

![Figure 5](image)

**Figure 5**
Bin1 mRNA expression detected by TaqMan quantitative real-time RT-PCR. **A.** In pigs 1, 3 and 4, bin1 mRNA was more abundant in the LD than in the psoas. In pig 2, bin1 expression in both muscles was similar. Pigs 1 to 3 were 22-week-old Berkshires, and pig 4 was a 7-week-old cross breed. **B.** Although ubiquitous, bin1 expression was by far most abundant in the brain and heart. Error bar = standard deviation.
previously found by quantitative real time RT-PCR that in at least 4 out of 6 pigs, of the same sex, age and breed as the one used in the microarray hybridisation, MyHC 2a and 2x were, in fact, more highly expressed in the psoas than in the LD [20]. The 3 fast MyHC isoforms found in Table 2 might have been the result of cross hybridisation by the relatively large amounts of MyHC 2b specific probe generated from the LD muscle. Secondly, in comparing the profiles of two normal physiological states, such as red and white muscles, large variation in normal gene expression between individual pigs could present a major problem [30]. This variation is mainly attributed to genetic differences that exist between individual pigs; even pigs of the same breed are not genetically the same. The use of pooled porcine mRNA samples could inadvertently increase the genetic variation within each experimental group of animals. Therefore, in the context of porcine red-white muscle microarray analysis, there may be no advantage in pooling mRNAs, derived from the same muscle of several pigs. On the other hand, the use of different muscles from the same pig could minimise the effects of environmental variation, which could exist between individuals. The use of inbred lines in laboratory rodents largely eliminates the problem of genetic variation between individuals of the same line. Hence, the use of labelled cDNA from pooled inbred individuals would enhance experimental reliability without increasing genetic variation. However, inbred pig strains are not widely available and limited to a few lines of mini-pigs. Whether the microarray probes were derived from an individual or pooled from a group of individuals, extensive validation, such as by quantitative PCR or Northern analysis, is necessary to demonstrate that the differential expression of a gene identified on a microarray is consistent in a wider context.

**Candidate genes for phenotype determination**

From the red-white muscle microarray results, a list of novel candidate regulatory genes that could influence muscle phenotype, such as hypertrophy, differentiation and isoform-specific expression, was identified. Most of the genes listed as unknown (Table 1 and 2) were found with major open-reading frames, suggesting that they code for protein products. Candidate regulatory genes with putative identities were based on homology comparison. Even for these genes, with the possible exception of bin1, their functional roles in skeletal muscle are largely unclear. CK2α1 subunit, smpx, and tyrosine kinase A6-related gene are particularly interesting. They were found to be more highly expressed in red muscle and are connected to the casein kinase 2 signalling pathway. CK2 is a serine/threonine kinase that has been implicated in cell growth and proliferation [31,32]. CK2-mediated phosphorylation of Myf-5, a member of a family of myogenic transcription factors, was reported to be required for Myf-5 activity [33]. Presently, the contribution of the CK2 signalling pathway to skeletal muscle function is not known. Its role in muscle phenotype determination requires further evaluation. The gene for heat shock 70 kD protein 8 (HSPA8) seemed to be upregulated in white muscle. Heat shock proteins are considered to be molecular chaperones and indicators of cellular stress [34]. The same gene was found to be upregulated in human hypertrophic cardiomyopathy [35]. It is not clear if this finding was related to cellular stress of myopathy or muscle hypertrophy.

### Table 3: Oligonucleotides and TaqMan fluorogenic probes

| Gene            | Primer | Sequence 5′→3′                      |
|-----------------|--------|-------------------------------------|
| β-actin S       | CCA GCA CCA TGA AGA TCA AGA TC    |
|                 | A      | ACA TCT GCT GGA AGG TGG ACA         |
|                 | Ps     | CCC CTC CCG AGC GCA AGT ACT CC     |
| GAPDH S         | AGG CTC GGG CTC ACT TGA A         |
|                 | A      | TGC CCA TCA CAA ACA TGG G          |
|                 | Ps     | AGC CAA AAG GGT CAT CAT TCG TGC CC |
| MyHC 2b S       | CAC TTT AAG TAG TTG TCT GCC TTG AG|
|                 | A      | GGC AGC AGG GCA CTA GAT GT         |
|                 | Ps     | TGC CAC GTT CAT CTG GTA ACA TAA GAGC|
| Bin1 S          | GCC AGC AAT GTG CAG AAG AA         |
|                 | A      | CAT CTG CCT TCC CCA GTT TC         |
|                 | Pa     | CCT TCT CCT GCG CGC GAG TGA G      |
| Novel gene (kc2725) S | TTT TTT CCA TTC CCT GGT TGA |
|                 | A      | AGG GAC CTT GTA AGC CAA CA         |
|                 | Ps     | CCA GTC TGG TGG CCT AGT CAT GCC C  |

S, sense; A, antisense, Ps, TaqMan sense probe; Pa, TaqMan antisense probe
Our results demonstrate the power of microarray analysis in identifying candidate genes that influence muscle phenotype. The challenge now is to confirm these associations and demonstrate how these genes are involved in the relevant muscle phenotypes.

**Methods**

**Construction of porcine skeletal muscle cDNA microarray**

Mass phagemid excision was performed on two porcine skeletal muscle λZAP-Express cDNA libraries (50-day foetal *longissimus dorsi* muscle and 3-day-old *gastrocnemius* muscle), developed in-house, according to manufacturer's protocol (Stratagene). Both libraries, each with a million primary plaques, had been extensively characterised by sequencing and screening, and were used to isolate several full-length cDNA clones, including MyHCs (6.0 kb in size) and transcription factors, such as NFAT2 and GATA2. One study, performed on adult human skeletal muscle, found that highly expressed genes represented only 9.1% of the transcript variation, whereas moderately and weakly expressed genes made up 27.5% and 63.5% of the variation, respectively [26]. In order to ensure that the clones chosen for the porcine microarray assembly represented as many different genes as possible, clones were screened by colony hybridisation to assess their relative abundance of expression. About 400 bacterial clones were picked onto each square agar plate (Bio-Assay dish, Nunc) for colony hybridisation with 32P-labelled cDNA probe, picked onto each square agar plate (Bio-Assay dish, Nunc) and transcription factors, such as NFAT2 and GATA2. One study, performed on adult human skeletal muscle, found that highly expressed genes represented only 9.1% of the transcript variation, whereas moderately and weakly expressed genes made up 27.5% and 63.5% of the variation, respectively [26]. In order to ensure that the clones chosen for the porcine microarray assembly represented as many different genes as possible, clones were screened by colony hybridisation to assess their relative abundance of expression. About 400 bacterial clones were picked onto each square agar plate (Bio-Assay dish, Nunc) for colony hybridisation with 32P-labelled cDNA probe, derived from mRNA extracted by oligo(dT)25 Dynabeads (Dynal) from the same muscle type and stage of development as that used for the library construction. Based on signal intensity on the autoradiograms, each clone was classified as a weakly, moderately or highly expressed gene [26]. Clones picked for microarray construction were judged to be lowly or moderately expressed. Miniprep plasmid DNA was prepared from selected clones using the QIAprep 96 Turbo kit (Qiagen).

Insert amplification of each pBK-CMV-based plasmid clone (Stratagene) was performed with T7 primer (5’-GTA ATA CGA CTC ACT ATA GGG C-3’) and T3 primer (5’-CGA AAT TAA CCC TCA GTC AAG GG-3’), using a HotStarTaq Master Mix kit (Qiagen). PCR was conducted in a 100 µl volume, using the equivalent of 0.1 µl of miniprep plasmid (about 15 ng), for 35 cycles at 55°C for 45 s, followed by 72°C for 3.5 min, and 94°C for 1 min, with an initial activation step of 95°C for 15 min. PCR products were purified with QIAquick 96 PCR Purification kit (Qiagen) and printed onto CMT-GAPS coated slides (Corning) using a Microgrid II arraying robot (BioRobotics, Cambridge, UK). Spotted DNA was immobilised by baking at 80°C for two hours. Each clone was printed in duplicate on each slide. Appropriate positive controls to help with orientation and vector controls were incorporated on the microarray.

**Red-white muscle microarray hybridisation**

Messenger RNA was extracted by oligo(dT)25 Dynabeads (Dynal) from the *longissimus dorsi* (LD), a white muscle, and the *psos*, a red muscle, of a 22-week-old pig (Berkshire breed). Dual Cy-dye labellings were performed with a CyScribe First-Strand cDNA labelling kit (Amersham), which incorporates the use of oligo(dt) primers and random nanomers. Labelled cDNAs were purified in AutoSeq G-50 columns (Amersham). In a typical Cy3 or Cy5 labelling reaction, using 0.5 to 1.0 µg mRNA, a final volume of 32 µl was obtained, which was used on two microarray slides. Prior to pre-hybridisation, microarray slides were denatured in distilled water by heating to 95°C for 2 min. Pre-hybridisation was then performed in a plastic 2-slide holder containing 15 ml of 3 × SSC (sodium chloride and sodium citrate), 2% bovine serum albumin (B4287, Sigma) and 0.1% SDS at 65°C for 20 min. After a brief rinse in distilled water at room temperature, the slides were dehydrated in absolute ethanol and centrifuged at 100 g for 2 min.

Labelled probe was not quantified. Equal volumes of Cy3 and Cy5 labelled probes (16 µl each) were mixed with 1.0 µl (8.0 µg/µl) of poly(dA) oligonucleotide (27–7836, Amersham), heated to 95°C for 2 min, and mixed with 33 µl of 2 × hybridisation buffer (GHB-200, Genpak). A microarray slide was layered with the hybridisation mixture, then covered with a 22 mm × 64 mm cover glass (BDH), to the exclusion of air bubbles, and placed in a hybridisation chamber (ArrayIt hybridisation cassette AHC-1) in a 45°C dry incubator for 24 h. Slides were washed once in 300 ml of 1 × SSC and 0.2% SDS for 10 min at 45°C, twice in 300 ml of 0.2 × SSC and 0.2% SDS for 10 min at 45°C, and twice in 300 ml of 0.1 × SSC for 10 min at 37°C. After drying in a centrifuge at 100 g for 2 min, scanning was performed with an Affymetrix 428 scanner. A total of seven microarray slides (replicates) were subjected to two-colour Cy-dye hybridisation and scanned for the red-white muscle experiment.

**Microarray expression analysis and clone identification**

Image analysis (grid generation and dye quantification) of scanned slides was performed with ImaGene v4.2 (Biodiscovery). Data mining was conducted with GeneSpring v4.2 (Silicon Genetics), in which 3 normalisation steps were performed: per spot, intensity-dependent, and per chip (slide). In per spot normalisation, after background subtraction, the fluorescent intensity (e.g. Cy3) of each clone was divided by its control or reference channel intensity (e.g. Cy5). Values of the control channel that fell below 10 were adjusted to 10 prior to taking the ratio between signal and control value. Intensity-dependent (non-linear or LOWESS) normalisation was applied to correct for artefacts caused by differential Cy3 and Cy5 dye incorporation and non-linear rates of fluorescence be-
between the two dyes. Per chip normalisation was used to take into account intensity variation across the entire slide, by dividing the signal strength of a clone on a slide by the 50th percentile signal of all of the measurements taken from the same slide. Additionally, the GeneSpring Global Error Model, based on the replicate measurement samples of all genes on the microarray was used to estimate differences between medians, the standard errors among the different clones, and the extreme tails of the distribution of differences. Differences among muscles in expression were standardised for different clones by dividing the difference between muscle medians by the square root of the common variance. Sequencing reactions with T7 and T3 primers were conducted with an ABI BigDye Terminator v3.0 PCR-based sequencing kit (Applied Biosystems). Performa DTR gel filtration cartridges (EDGE Biosystems) were used for purification of PCR products prior to loading onto the ABI PRISM 3100 Genetic Analyzer. All sequence data were subjected to BLAST (Basic Local Alignment Search Tool) searches for gene identification by sequence similarity.

Quantitative real-time RT-PCR
Quantitative real-time RT-PCR was performed using Taq-Man (Applied Biosystems) on five selected porcine genes: β-actin, GAPDH, MyHC 2b, bin1, and a novel gene (kc2725) (Table 3). The protocol, based on the use of the relative standard curve, was as previously described [20]. In the relative standard curve method of quantification, comparisons of relative expression should only be made between samples of the same gene. It is not appropriate to compare expression levels between two different genes. A reference cDNA panel, comprising a number of different tissue templates (LD muscle, psoas, heart, uterus, brain, liver and spleen) of 7-week-old pig, along with LD and psoas cDNAs, from three additional 22-week-old Berkshire pigs, were used to evaluate the quantitative distribution of selected genes.

Authors’ contributions
QB generated the microarray clones. CM, NdC, MJM and KCC helped with the validation and analysis of the microarray. DS and GE were responsible for the production of the microarray and advised on its experimental use. KCC conceived the study and was responsible for writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements
This work was supported by a scholarship from the University of Glasgow to QB, by Sygen International and by the BBSRC.

References
1. Pette D and Staron RS Cellular and molecular diversities of mammalian skeletal fibers Rev Physiol Biochem Pharmacol 1990, 116:2-76
2. Schiaffino S and Reggiani C Molecular diversity of myofibrillar proteins: Gene regulation and functional significance Physiol Rev 1996, 76:371-423
3. Schiaffino S and Reggiani C Myosin isoforms in mammalian skeletal muscle J Appl Physiol 1994, 77:493-501
4. Weiss A and Leinwand LA The mammalian myosin heavy chain gene family Annu Rev Cell Dev Biol 1996, 12:417-439
5. Chang KC, Fernandez K Developmental expression and 5’ end cDNA cloning of the porcine 2x and 2b myosin heavy chain genes DNA Cell Biol 1997, 16:1429-1437
6. Chang KC, Fernandez K and Dauncey MJ Molecular characterisation of a developmentally regulated porcine skeletal myosin heavy chain gene and its 5’ regulatory region J Cell Sci 1995, 108:1779-1789
7. Chang KC, Fernandez K and Goldspink G In vivo expression and molecular characterisation of the porcine slow-myosin heavy chain gene J Cell Sci 1993, 106:331-341
8. Soares RE, Brooks L and Dauncey MJ Muscle fibre type and meat quality Meat Sci 1998, 49:521-529
9. Karlsson AH, Klont RE and Fernandez X Skeletal muscle fibres as factors for pork quality Livest Prod Sci 1999, 60:255-269
10. Essén-Gustavsson B Muscle-fibre characteristics in pigs and relationships to meat-quality parameters review Pork quality: genetic and metabolic factors (Edited by: Puolanne and Demeyer(Dl) Walfingford UK, CAB International 1993, 140-159
11. Hughes SM and Schiaffino S Control of muscle fibre size: a crucial factor in ageing Acta Physiol Scand 1999, 167:307-312
12. Larsson L and Ramamurthy B Aging-related changes in skeletal muscle Drugs Aging 2000, 4:303-316
13. Duggan D, Bittner M, Chen Y, Meltzer P and Trent JM Expression profiling using cDNA microarrays Nat Genet 1999, 21:10-14
14. Blanquet PR Casein kinase 2 as a potentially important enzyme in the nervous system Prog Neurobiol 2000, 60:211-246
15. Faust M and Montenarh M Subcellular localization of protein kinase CK2. A key to its function? Cell Tissue Res 2000, 301:329-340
16. Kemp TJ, Sadusky TJ, Simon M, Brown R, Eastwood M, Sassoon DA and Coulton GR Identification of a novel stretch-responsive skeletal muscle gene (Smqpx) Genomics 2001, 72:260-271
17. Rohwer A, Kittstein W, Marks F and Gschwendt M Cloning, expression and characterization of an A6-related protein Eur J Biochem 1999, 263:518-525
18. Sabatini D, Elliott KJ, Wechsler-Ray and Prendergast GC Bin1 is a novel Muc-interacting protein with features of a tumour suppressor Nat Genet 1996, 14:69-77
19. Wechsler-Ray and Prendergast GC A role for the putative tumor suppressor bin1 in muscle cell differentiation Mol Cell Biol 1998, 18:566-575
20. da Costa N, Blackley R, Alzuherri H and Chang KC Quantifying the temporo-spatial expression of porcine postnatal skeletal myosin heavy chain genes J Histochem Cytochem 2002, 50:353-364
21. Tsutsui K, Meeda Y, Seki S and Tokunaga A cDNA cloning of a novel amphibins isoform and tissue-specific expression of its multiple variants Biochem Biophys Res Commun 1997, 236:178-183
22. Rink A, Sanschi M and Beattie CW Normalized cDNA libraries from a porcine model of orthopedic implant-associated infection Mammal Genome 2002, 13:198-205
23. Soares MB, Biondino MF, Jelene P, Su L, Lawton L and Efstratiadis A Construction and characterization of a normalized cDNA library Proc Natl Acad Sci U S A 1994, 91:2289-2322
24. von Stein OD Isolation of differentially expressed genes through subtractive suppression hybridization Genomics protocols (Edited by: starkey and Elaswarapu) Totowa, Hum Press Inc 2001, 263-278
25. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Smith HO, Yandell M, et al. The sequence of the human genome Science 2001, 291:1304-1351
26. Yandell M, Brandon Rhonda, Cargill Michele, Chandramouliswaran Ishwar, Karvitz Saul, Levy Samuel, Mobbary Clark, Reinert Kurt, Remington Karin, Abu-Threideh Jane, Beasley Ellen, Biddick Kendra, Bonazzi Vivien, Brandon Rhonda, Cargill Michele, Chandramouliswaran Ishwar,
29. Nadon R and Shoemaker J
30. Chang KC, da Costa N, Blackley R, Southwood O, Evans G, Plastow B
31. Guerra B and Issinger OG
32. Tawfic S, Yu S, Wang H, Faust R, Davis A and Ahmed K
33. Winter B, Kautzner I, Issinger OG and Arnold HH
34. Liu YF and Steenacker JM
35. Lim DS, Roberts R and Marian AJ

BMC Genomics 2003, 4
http://www.biomedcentral.com/1471-2164/4/8

Bortoluzzi S, d’Alessi F, Romualdi C and Danieli GA

Page 13 of 13
(page number not for citation purposes)