

# CorrDB: A livestock animal genetic / phenotypic trait correlation database

Zhi-Liang Hu<sup>1</sup>, Dinesh Kumar<sup>2</sup> and James M. Reecy<sup>1</sup>

1. Department of Animal Science and Center for Integrated Animal Genomics, Iowa State University, 2255 Kildee Hall, Ames, IA 50011

2. Genes and Genetic Resources Molecular Analysis Lab, National Bureau of Animal Genetic Resources, India

## Abstract

Genetic and phenotypic trait correlation studies have been conducted for almost 70 years. A phenotypic correlation is the degree to which two traits co-vary among individuals in a population; A genetic correlation is the proportion of shared variances between two traits due to genetic causes. In the genomics era, while gene expression, genetic association, and networked analysis provide un-precedent means to decode the genetic basis of complex phenotypes, it will be important to be cognizant of possible effects on other traits should genetic progress be made in another. Furthermore, a genetic/phenotypic trait correlation database should help to facilitate genetic network analysis and systems biology studies. The CorrDB (<http://www.animalgenome.org/CorrDB/>) was designed to collect all published livestock genetic/phenotypic correlation data published to date. The database was developed on MySQL DBMS. The user and curator interfaces were programmed in Perl CGI and PHP. We have emphasized capabilities to dynamically link to external databases for genes and traits such that tools may be built for potential data integration and analysis. Currently, this database has an initial collection of 3,635 correlations on 276 cattle traits, relating to meat production, milk production, growth, health, etc. This correlation data was curated from 143 journal papers. Our long-term goal is to make it part of the integrated functional genomics resources for all livestock animals.

## Introduction

Trait correlations are common phenomena in biology. It has been a common practice among animal breeders to consider trait correlations to either improve correlated traits simultaneously or to reduce undesirable side effects when improving only one of the correlated traits. Since Hazel (1) showed in 1943 how the genetic correlation coefficient measures the association between two traits, and demonstrated the application of it as a predictor in selection indexes, many studies have reported genetic and phenotypic correlation estimates between traits in livestock animals, along with other genetic parameters such as heritability and variances measured under various experimental conditions. These data are useful not only in animal breeding practice, but also in studies dissecting the underlying genes and genetic mechanisms responsible for quantitative trait variations between animals.

For example, pleiotropy or close linkage are the two major reasons for genetic trait correlations and are often confounded at the level of quantitative trait loci or genes. With the progress of genetic and genomic approaches, discrimination of intragenic linkage from true pleiotropy is increasingly possible. This will help to improve the understanding of the nature of trait correlations and underlying genetic mechanisms, which will have substantial impact on the progress in making breeding strategies.

The purpose of constructing an animal trait correlation database is to make useful of historically abundant phenotypic data in current genomic studies where it is becoming more feasible for networked analysis of correlations between traits, quantitative trait loci (QTL), coding gene expression and interactions, single nucleotide polymorphisms (SNP) and other types of genetic variations.

## Results

The CorrDB was developed on MySQL DBMS, and the user and curator web interfaces were programmed in Perl CGI and PHP. Currently, this database contains an initial collection of 3,635 correlation data from 143 journal papers on 276 cattle traits, relating to meat production, milk production, growth, health, and other traits (Figure 1).

1. The user interface provides four data view options: Matrix View (default; Figure 2), Graphic View (Figure 3), Table View (Figure 5), and plaintext data download format. A search box and results filters are implemented as a tool bar on the top of the pages to serve as a data exploration tool.
2. The Graphic View was implemented with free graph-drawing software called Graphviz (2). The utility of the graph tool is demonstrated by examples in Figures 3 and 4, where groups of closely related traits are clearly seen.
3. Dynamic hyperlinks for further exploration of traits and references on the queried data are currently available only in Table View (Table 5; the hyperlinks on other formats will also be available soon). Note that links to QTLdb will come soon, once successful mappings of the traits are made.

Figure 4.

Examples of larger picture views of correlation data "Graphic View". These examples demonstrate its utility as a guide for groups of closely related traits.

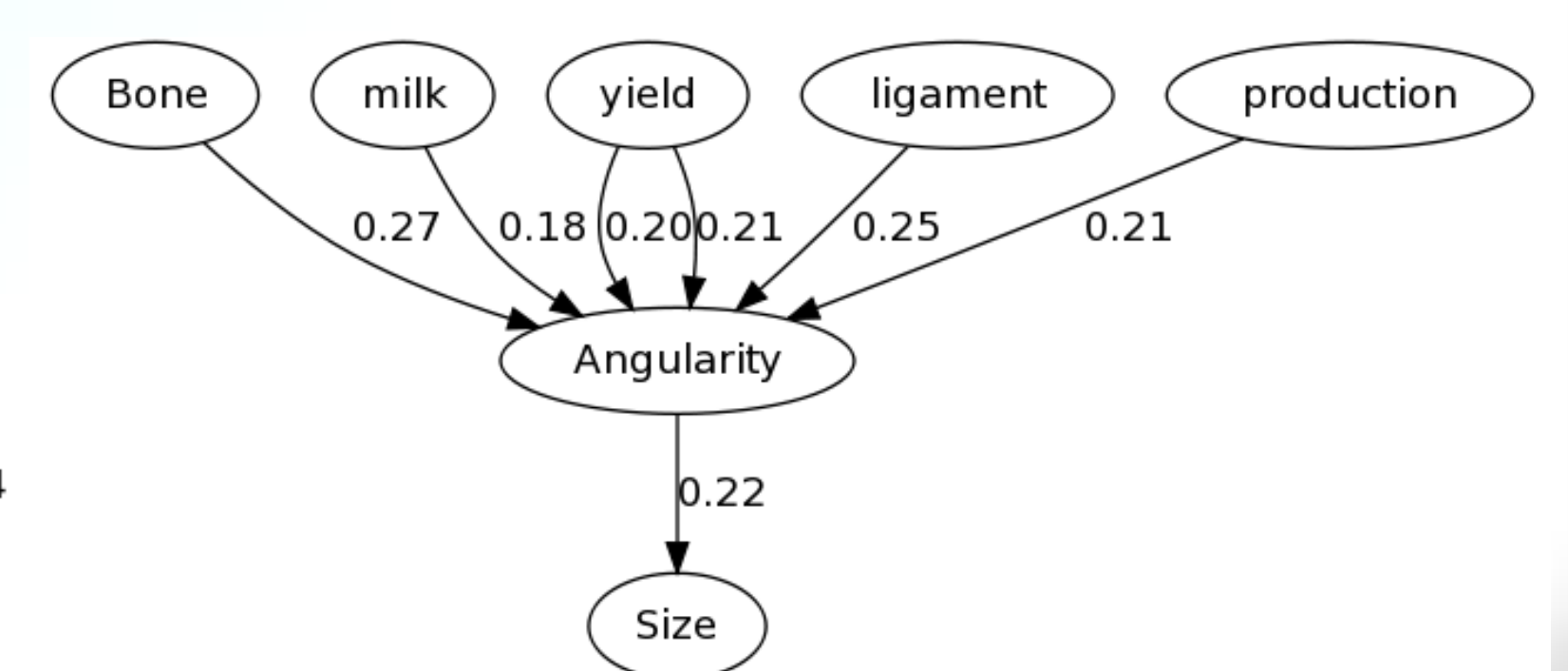
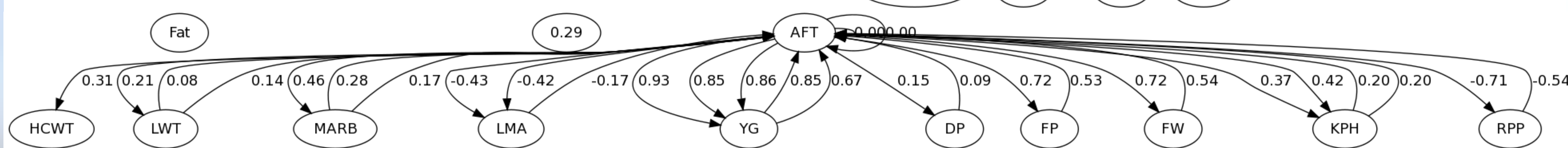


Figure 1.

The CorrDB front page showing the database content summary, data access points, and our open policy for new data curation. One of the principals we tried to follow is to allow users get to the data in as few steps as possible, and in ways as simple as possible.

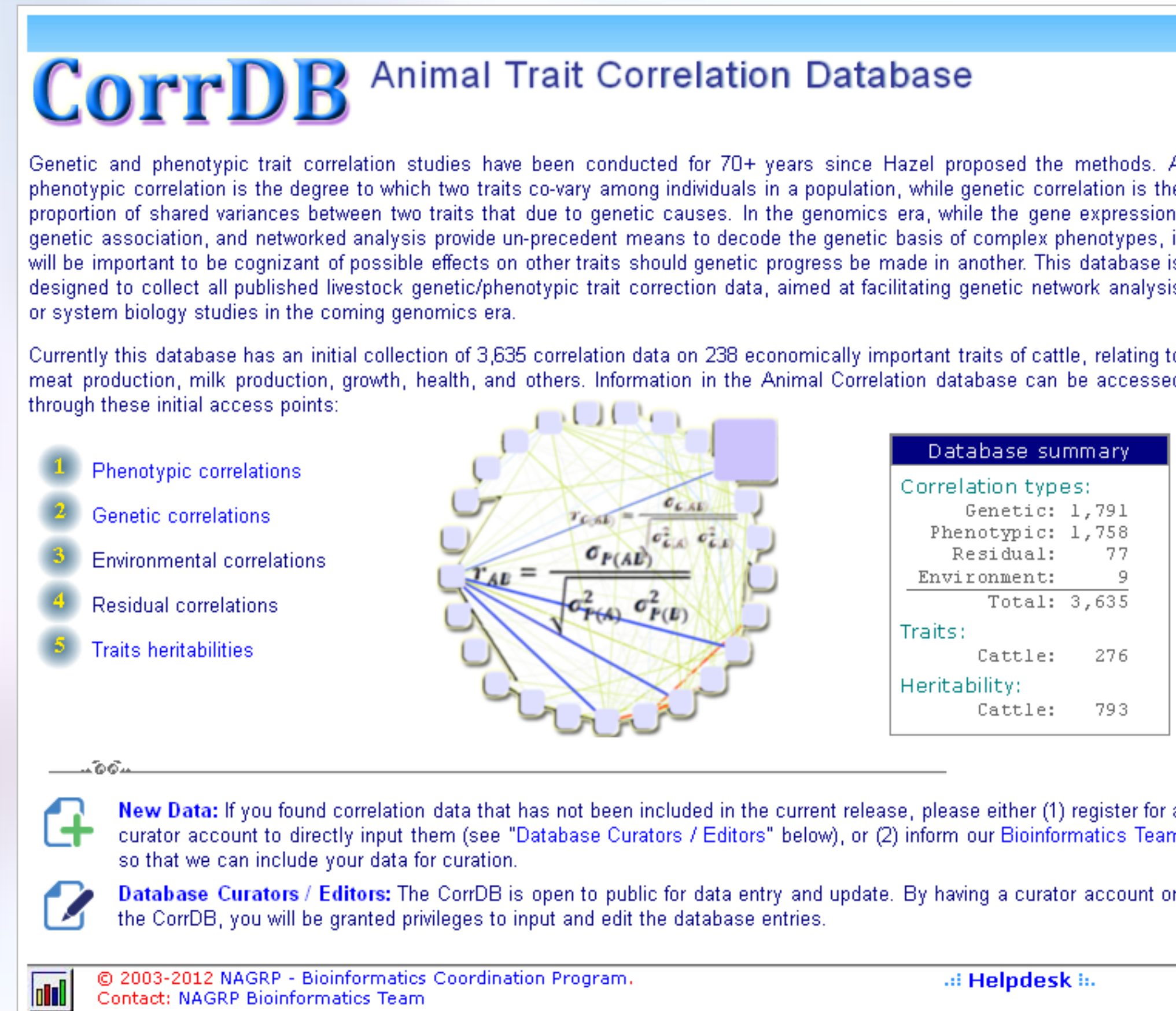


Figure 2.

A screen shot of the CorrDB data "Matrix View" where pairs of correlated traits are placed in a two-dimensional table. Note that the same data set can also be viewed in "Table View" and "Graphic View", and can be downloaded in plain text. The tool bar on the top can be used to search for data by traits, by correlation types, and by species.

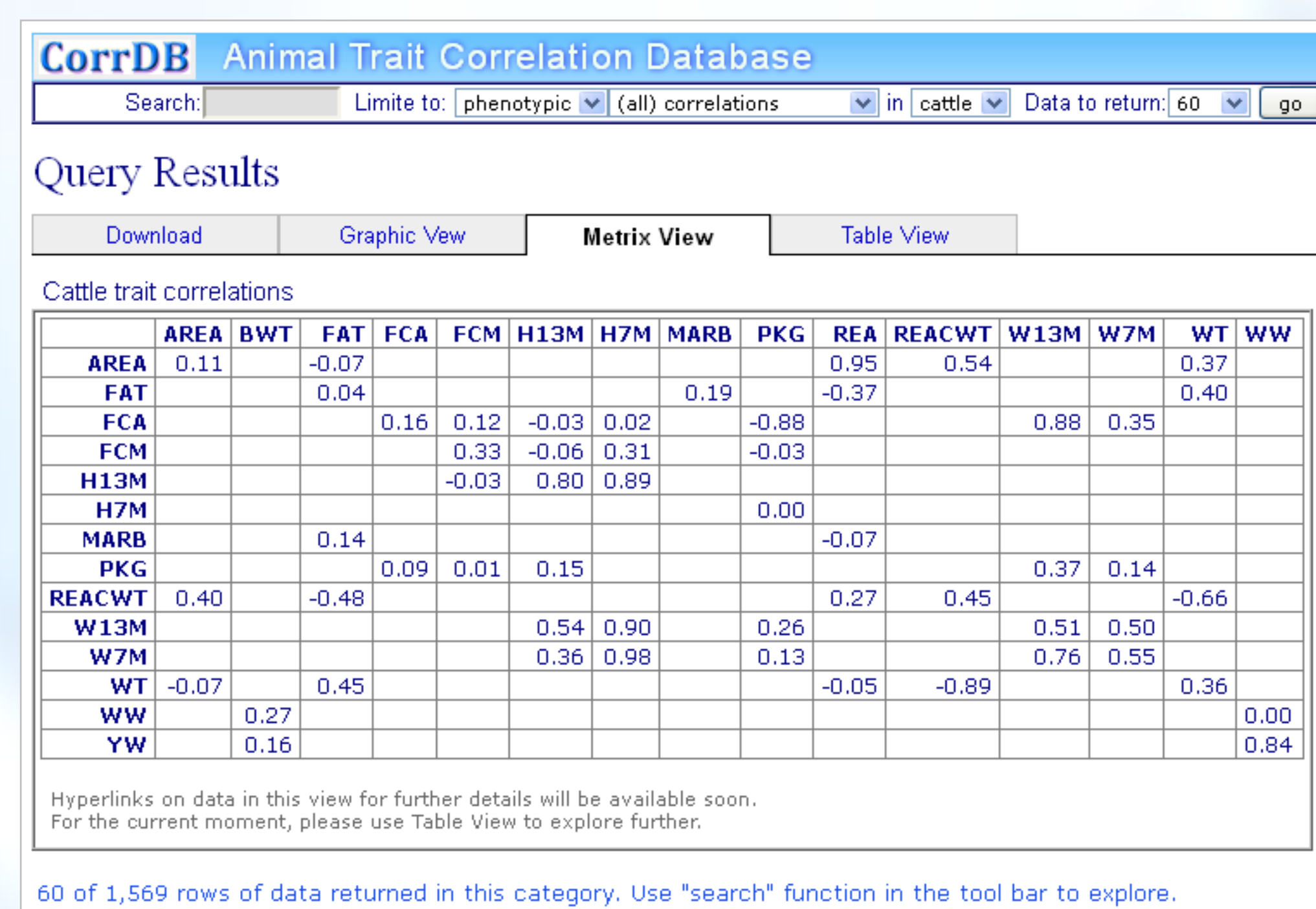


Figure 3.

A screen shot of the CorrDB data "Graphic View" of the same set of data (as seen in Figure 2). This view helps to give users an idea how the correlated traits are aggregated in order to assist with further data analysis.

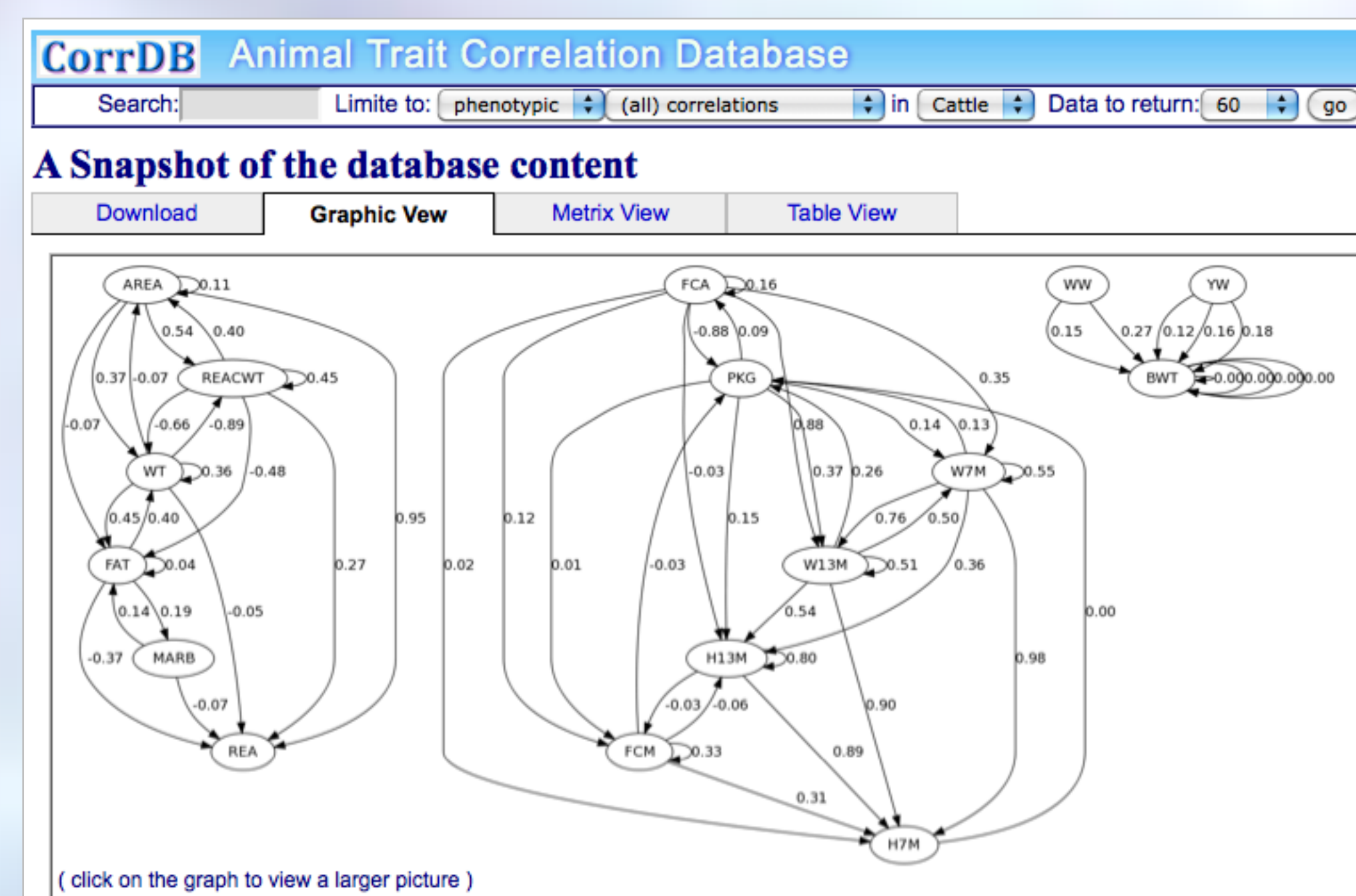
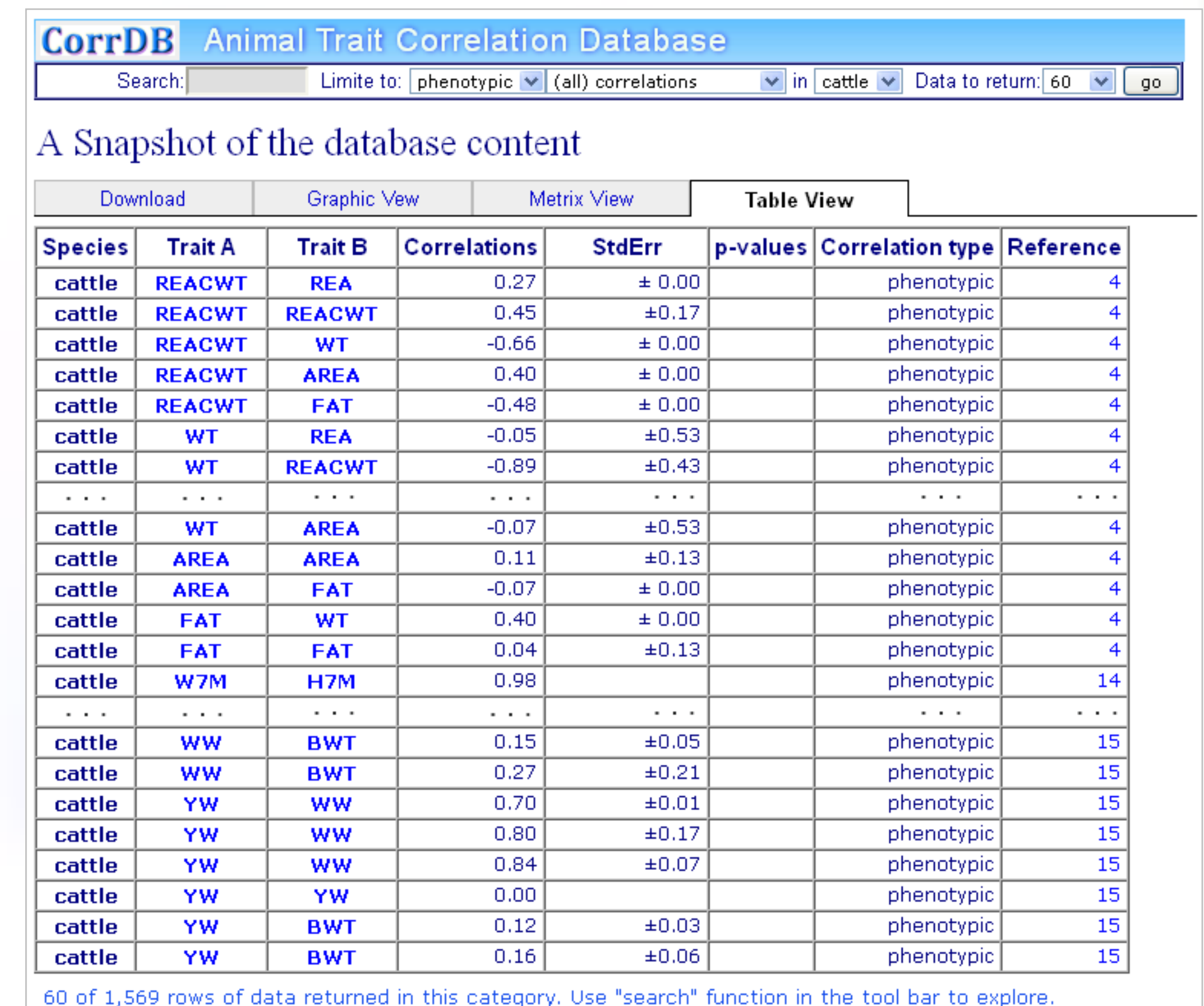


Figure 5.

A screen shot of the CorrDB data "Table View" of the CorrDB data display, where users can explore for further information such as trait details, referenced publications, etc. Note the data summary given on the bottom of the page and the utility of a tool bar on the top to allow users to explore the content of the database. The dotted cells represent omitted lines (in order to fit the limited display space here).



## Results (continued)

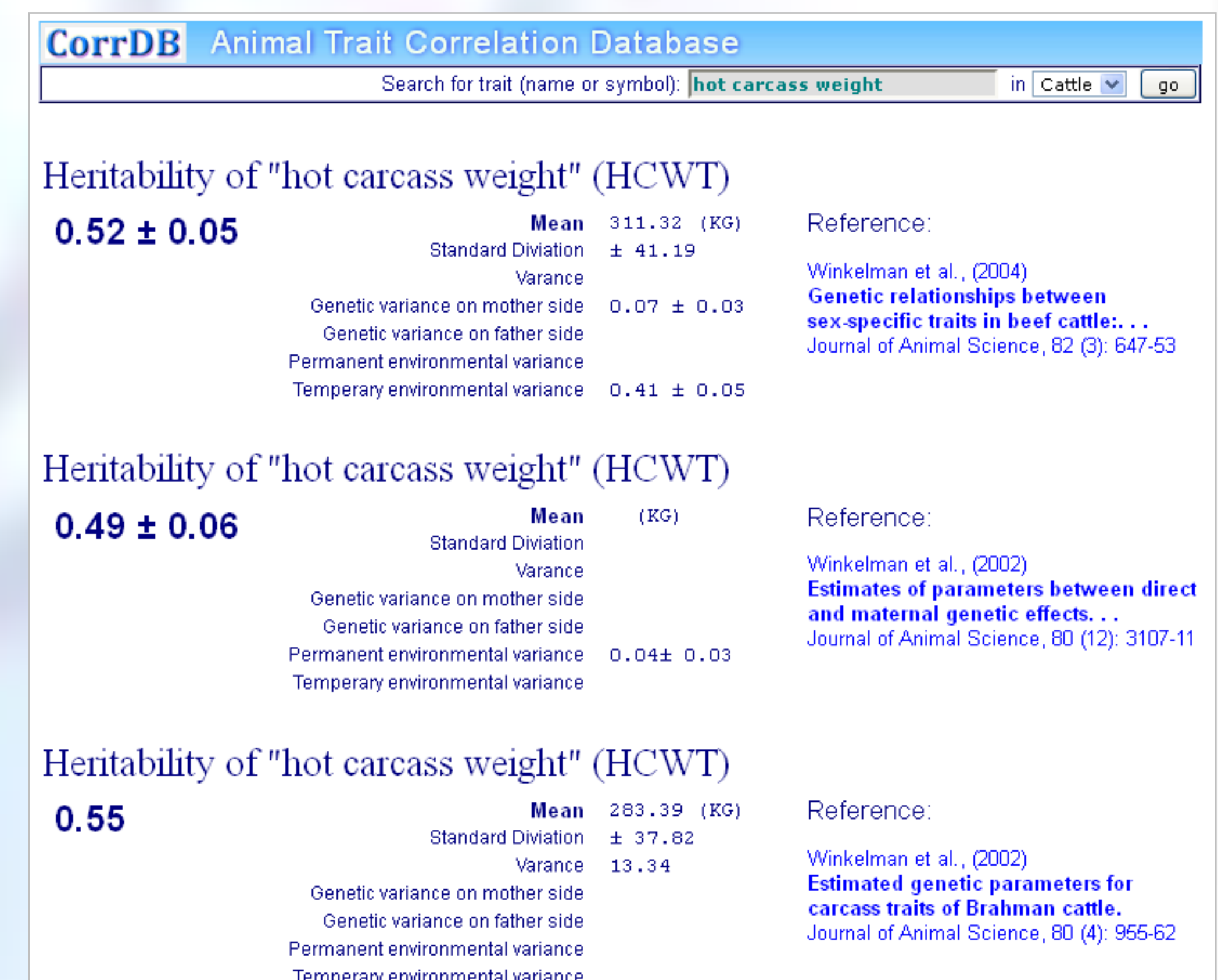
4. Trait heritability data is also made available in CorrDB since the trait heritability and correlation are often reported together. In Figure 6 is shown heritability for "hot carcass weight" derived from different experiments, with trait measurements and related information listed for reference.

The CorrDB was designed to contain trait attribute information such as trait descriptions, measurement methods, experimental conditions, and trait ontology mappings. Future developments of the CorrDB may also be annotated with necessary genomics information in order to make it part of the integrated functional genomics database resources by linking out to other databases.

Currently the web portal is still under active development. More web functions will be available soon.

Figure 6.

A snapshot of the heritability data display page, where heritability data from different experiments are listed, and trait measurements are given for reference.



## References

1. Hazel L. N. [1943]. The genetic basis for constructing selection indexes. Genetics 28, 476-90.
2. Ellson, J., Gansner, E.R., Koutsofios, E., North, S.C., Woodhul, G. (2003). Graphviz and dynagraph - static and dynamic graph drawing tools. In: GRAPH DRAWING SOFTWARE (2003), pp. 127-148.