

## PEDIGREE ANALYSIS by Byron Rogers

# FINDING

## *Tregonwell's Natural Barb mare*



*Black Caviar*

(Bel Esprit-Helsing by Desert Sun (GB), family 1-p)

ALL modern thoroughbreds trace back in male line to one of three horses of Eastern origin, the Darley Arabian, the Godolphin Arabian, and the Byerley Turk, and through only three descendants, respectively Eclipse, Matchem and Herod. In direct female line, the breed traces to a larger group of foundation matriarchs, including both foundation stock native to the British Isles known for their speed, such as the Galloway, from the north of England, and the Irish Hobby, with imported strains from Asia.

In 1791, James Weatherby published an Introduction to a General Stud Book, an attempt to collect pedigrees of current and previous racehorses. The first volume of the General Stud Book was published two years later, with notable revisions taking place in 1803, 1808, 1827, 1859 and 1891. Subsequently, to become a registered thoroughbred, a horse had to trace its ancestry to horses already registered in the General Stud Book. The records of the General Stud Book, incorporating records from private stud books that pre-existed the formation of a register, stretch back more than 300 years, and ensure the thoroughbred is one of the most richly documented species on earth. These records are now maintained in the General Stud Book of England (GSB) and other breed registries around the world in a well-organised fashion.

### Bruce Lowe Family Numbers

THE lineage of thoroughbred racehorses has, for a significant period of time, held importance in the racing and breeding industry with unproven stock being valued in many cases on their immediate maternal lineage. Indeed the catalogue page highlights the female family of the given yearling or mare over other parts of the pedigree which may have just as much influence on the outcome.

Traditional thoroughbred pedigree analysis has classified the maternal descent of the breed in terms of 'Family Numbers', with these numbers often being carried in catalogue pages up until the mid-1980s. The family numbers commonly used to designate various thoroughbred female families were popularised by C. Bruce Lowe, an Australian pedigree researcher at the end of the nineteenth century and flourished at the beginning of the next century with a posthumous publication of his works.

Lowe had traced back the pedigrees of the complete list of winners of the three oldest English classic races, the St Leger Stakes, Epsom Derby Stakes and Epsom

Oaks, grouping them by direct lines of tail female descent, from dam to grand-dam and on back until the family was no longer traceable in the GSB.

Families were then assigned a number by Lowe based on the total number of classic winners descended from the family at that time. The family descending from Tregonwell's Natural Barb Mare, was designated "#1 Family", the Burton Barb Mare, whose descendants had produced the second highest number of classic winners, designated "#2 Family", and so on. The resulting 43 numbered families became the core of his study, and many thoroughbred breeders still use his family numbers as a convenient way to categorise thoroughbred families to this day.

Additional works such as The Family Table of Racehorses, also commonly referred to as the Bobinski Tables, which were the first to designate different branches of specific families with the addition of a letter such as 1-s (for Web, a 1808 mare who descended from Tregonwell's Natural Barb) or 2-f (for the Hyacinthus mare, an 1804 descendant of Burton Barb mare), with the most recent updates known as the Toru-Shirai tables. It must be understood

## PEDIGREE ANALYSIS

that the designation of these branches of families came without scientific basis, but arbitrarily based upon a matrilineal descendant of the taproot mare having a number of offspring who were distinguished as performers or producers.

### #1 Family – Tregonwell's Natural Barb

THE exact origin and ownership of Tregonwell's Natural Barb is not clear as her stud record pre-dates both the first edition of the GSB, and many of the private records which were the original source material for the GSB. A 1744 advertisement for the stallion Merry Andrew in the Dublin Journal called her "a natural Barb mare of Mr Tregonwell's", while a pedigree for the well known stallion Whitefoot (1719) in the Dutton Stud records make reference to his ancestor being "a Natural Barb Mare of Mr Tregonwell's a Noted Breeder of the South".

It is believed Tregonwell's Natural Barb was born around 1657 and lived until around 1670. Records from the first edition of the GSB do reflect that she had at least three known foals: a filly by Place's White Turk; a filly by Restive; and a colt named Rockwood, whose sire was not stated, although the various forms of reference that were made to associate these foals to Tregonwell's Natural Barb do cast doubt as to the accuracy of this breeding record. By Stud Book records, today's members of the #1 Family descend from Bonny Lass (1723), a bay mare bred by Sir William Ramsden and belonging to the Duke of Bolton. Very little is known about her ability as a racehorse, but at stud she produced Merry Andrew (1730) for the Duke of Bolton and later for her new owner, a Mr Panton, the Partner Mare (1735) from whom most, if not all, the #1 Family descends.

An examination of the 4181 yearlings offered at the 2014 September Yearling Sales (Keeneland) by matching all the yearlings offered at that sale to their Bruce Lowe family numbers indicates that the "#1 Family" by Stud Book records represents approximately 15% of the yearlings catalogued at that sale and is by Bruce Lowe numbers by far the most populous.

### Mitochondrial DNA – where science steps in

WHILE the thoroughbred industry is yet to fully embrace it, science has caught up with and well passed Bruce Lowe numbers. The domestic equine mitochondrial genome was first sequenced by Xu and Arnason in 1994 and was found to contain approximately 16,660 nucleotides. Common to all mammalian species, Mitochondrial DNA (mtDNA) is inherited exclusively from the mother and mutations regularly occur, often in adaptation for environmental conditions. Sets of mutations shared by a large number of individuals can be used to define maternally linked populations known as haplogroups, and sub-populations known as haplotypes. This has made mtDNA a very useful tool for studying the evolution of, and classifying, various species including the horse in a far more accurate way than any stud book or record keeping system could.

Using mtDNA mutations as markers for familial segregation, Dr Alessandro Achilli and colleagues from the University of Perugia have defined mutational markers of haplogroups and haplotypes for the entire equine breed, making it possible for classification of the thoroughbred not only in relation to within the breed, but also when compared to other equine breeds. In 2002, Dr Emmeline Hill and colleagues, were the first to publish a scientific paper on maternal classification of the thoroughbred via mtDNA analysis,



*Pierro*  
(Lonhro-Miss Right Note (IRE) by Daylami, family 1-t)

and by analysing the mtDNA of 100 thoroughbreds that were supposed to represent 19 female families according to General Stud Book records.

Using just a small fraction, 381 bases of the 16,600 mtDNA nucleotides, they identified 17 haplotypes and concluded that there were significant differences between General Stud Book records and mitochondrial haplotypes in the breed. A few years later in a larger analysis using most of the mitochondrial genome, Drs Harrison and Turrion-Gomez also identified a total of 17 haplotypes in the thoroughbred, but observed that variation occurred across the entire mitochondrial genome.

Further studies have more recently been conducted by Dr Mim Bower and colleagues from the University of Cambridge, and while still only considering a small portion of the total mtDNA they established some significant discrepancies between the mtDNA and General Stud Book records. As an attempt to clarify the exact extent of the errors in the stud book, last year with Dr Gus Cothran and Anas Khanshour from Texas A&M University, I undertook a private study where we sequenced and analysed two regions of the mtDNA of 90 thoroughbreds, many believed to be representing the #1 Family, and compared these sequences against a reference sequence and sequences from previous thoroughbred mtDNA studies and other breeds. The sequences are now publicly available for reference in Genbank, an online database of DNA sequences for scientific study.

Using Achilli's reference genome and haplogroups we established that like previous studies by Dr Hill and colleagues and Dr Bower and colleagues, our sequences from the #1 family placed this family broadly into the "N" haplogroup which is commonly shared with other equine breeds including sequences from the English Shire, Andalusian, Saddlebred, Zhongdian and Exmoor. Drilling down from the haplogroup to the haplotype level by using further mutations to separate out the #1 family from these other breeds, we found the

sequences place the "#1 Family" in agreement with a subdivision/haplotype of the "N" haplogroup with identical mutations to an Andalusian sequence.

This subdivision/haplotype within the "N" haplogroup is unique with five distinct markers located along the mitochondrial genome which separate it from all other sequences within the thoroughbred breed. The fact that the #1 family samples we collected were from horses in five different countries, yet the sequences showed they clearly belong to the same mtDNA haplotype speaks volumes for the quality of record keeping over the past 300 years. The 90 new mtDNA sequences have also allowed us to clear up some misconceptions with the history of the #1 family from previous mtDNA papers, clarify some common ancestors, and confirm some errors that currently abound in breed classification.

### The case of Bend Or (1877)

IN 2012 Dr Bower and colleagues completed a multidisciplinary study of historic thoroughbred horses that included the 1880 English Derby winner Bend Or whose win was shrouded in some controversy as the owners of Bend Or were accused of swapping him, by accident or otherwise, with another horse Tadcaster, whose maternal pedigree was more prestigious at the time.

By Stud Book records, Bend Or was by Doncaster out of the mare Rouge Rose, subsequent foundation mare of the 1-k branch of the #1 family. mtDNA collected by Dr Bower from Bend Or's skeleton which resides at the Natural History Museum in London however, indicated that his mtDNA sequence did not match with other members of the #1 family, rather it was a haplotype shared by the #2 family, from which by GSB records Tadcaster came from and thus the skeleton known as Bend Or is most probably that of Tadcaster.

In our own analysis of the sequence and other sequences from both the #1 and #2 family, Bend Or's

## PEDIGREE ANALYSIS

sample places him in the “L” haplogroup, similar to other samples from the #2 Family affirming Bower’s findings. This case of the skeleton of Bend Or and its mitochondrial haplotype not agreeing with stud book records is a good example of how prior to DNA parentage verification arriving in the mid 1980s, errors could occasionally creep into even the most meticulously kept stud book records. The irony that Rouge Rose was assigned by the Bobinski tables as a founder mare for the 1-k branch mostly due to the efforts of “Bend Or” on the racetrack and at stud, where he was a multiple leading sire, should not be lost on the reader.

### Maid of the Glen (1858)

BOTH of the studies completed by Dr Hill, et al and Dr Bower, et al described a discrepancy from Stud Book records for the branch of Maid of the Glen, the founder of the 1-u branch. In our study two samples with common ancestor of Furze Bush (1891) were sequenced with contrasting results. One sample, from a European bred and raced stallion tracing to Whinbloom (1901 daughter of Furze Bush) found agreement with Dr Hill and Dr Bower with the sample sharing the “L” haplogroup.

However, another sample, a branch of the 1-u tracing to another daughter of Furze Bush in Veldt (1905) that is more popular and was sourced in North America, did not agree with the studies by Dr Hill and Dr Bower, rather this sequence agreed with the “N” haplogroup of the #1 family. It is hard to be exactly certain when this discrepancy occurred other than to say that based on the three studies completed on this branch of the #1 family, some or all of the branch of 1-u via Whinbloom (1901) does not share the same mitochondrial haplotype as other members of the 1-u family.

### Web (1808) and La Troienne (1926)

DR BOWER and colleagues had suggested errors existed in the stud book between the #1 Family lineage and the #2 Family lineage and could be traced to Web (1-s). Four samples from our study tracing directly to Web through lines exclusive of any of the branch (1-t, 1-u, 1-w, 1-x) taproots under her were obtained. Two of the samples had a Most Common Recent Ancestor (MCRA) of Absurdity (1903) (12 and 11 generations respectively) the third sample shared a MCRA with the first two, that of Absurdity’s dam Paradoxical (1891), while the fourth, a rarer branch of the family, shared a MCRA with the other three samples at Filagree (1815), a daughter of Web by Stud Book records.

A further 16 samples in our study traced to the imported mare La Troienne, one of the most famous mares of the American Stud Book who shares an MCRA of Paradoxical with three of the samples described above and Filagree with the fourth sample. The 16 samples obtained traced to seven different daughters of La Troienne and all 16 through different granddaughters of the mare giving wide coverage of this commercially popular branch. All 20 samples shared the “N” markers of the #1 family indicating that they all belonged to the same family.

It is apparent Bower and colleagues misplaced the maternal lineage of a discrepancy they found in a sample tracing to 1-p (Hilarity) as occurring with Web, when in fact that branch on stud book records traces to a sister to Web (that of Wire b.1811). Similarly a discrepancy found by Bower in the 1-n branch of Chelandry (1894) attributed to Web seems to be a misread by that study of stud book records, as Chelandry traces matrilineal to Pawn (1808) who is a sister to Web’s dam in Penelope (1798).



*Super Saver (USA)*

*(Maria’s Mon-Supercharger by A.P. Indy, family 1-s)*

### Chelandry (1894)

BOWER and colleagues also identified errors in their study in the branch of Chelandry (1-n) in three samples. For our study, seven samples matrilineal tracing to Chelandry were obtained with a sample tracing to each of her daughters in Bobolink, Chelys, Skyscraper and Yippingale and four samples for Popinjay (via four distinct daughters of that mare in Gay Bird, Pompadour, Poppingal and Prattle). Chelandry is a granddaughter of Paraffin (1-l), a further five samples were obtained tracing to unique daughters and granddaughters of Paraffin. All 16 samples in this study shared the “N” haplogroups markers. As Bower and colleagues study contained three samples that did not agree with the “N” haplogroup, rather the “L” haplogroup, yet our 16 samples did, it is probable that the three samples obtained by Bower, et al represent recent stud book anomalies.

### Family #16

DR HILL and colleagues found that one of the eight horses analysed from what is known as the #16 Family, that tracing to founder mare Stripling’s Dam, showed the “N” founder haplotype, suggesting commonality with the #1 family, with the rest of her samples from the “L” haplogroup. A descendant of the mare Lady Alice, an 1855 filly by Chanticleer out of the mare Agnes, was the horse whose genetic sequence displayed the anomaly in that study. To verify this we obtained two modern day samples tracing to Lady Alice, 11 samples from branches descending from Lady Alice’s half-sister Miss Agnes and two samples descending from Caroline (1836), a half-sister to Lady Alice’s grand-dam.

Analysis of the sequences revealed this female line has some significant deep-rooted stud book discrepancies in it. Of the two samples tracing to Lady Alice, one agreed with the “L” haplogroup found in seven of Hill and colleagues samples from this family. The other sample found agreement with the “I” haplogroup,

not the “N” haplogroup as was expected based on Hill’s study. Furthermore, of the 11 samples descending from Miss Agnes, a half-sister to Lady Alice, one sample, that tracing to Wild Agnes (1862), founder of the 16-d branch of this family displayed the “I” haplogroup. These studies show there may be significant historical Stud Book errors in the #16 family.

### American Family #1 (Janus Mare)

BOWER et al suggested that the Janus Mare No 1, annotated as the American #1 or A1 Family, could well be part of the #1 Family via sequence agreement. The two samples in this study that were sequenced did not agree with Dr Bower’s sequence of the A1 Family and more startlingly did not find themselves in agreement with the markers of the “N” haplogroup associated with the #1 Family at all. Rather these two samples aligned with the “G” haplogroup. As the two samples selected for this study had a MCRA of Ballet (1871) and were 14 and 13 generations respectively to that MCRA it is more likely that the A1 Family is from the “G” haplogroup and that the single sequence used in Dr Bower to ascribe joint matrilineal descent of the #1 and A1 families is indicative of a Stud Book error in their sample. Further sequences from this family would clarify this situation.

### American Family #4 and Family #25

USING a small fraction of the mtDNA sequence Dr Bower and colleagues described that the American A4 Family (Medley Mare No 2) is separated from the #1 Family by a single mutation. The two samples from our study agreed with Dr Bower and colleagues sequence, however larger sequences for the same samples from A4 Family indicate that it is a separate haplotype of the “N” haplogroup with a mutation outside the area tested by Dr Bower and colleagues.

The A4 Family shares a “G” mutation common with the English Shire and Saddlebred and lacks the “T” mutation indicative of the #1 Family as described above. Thus the A4 Family could not be, as Dr Bower proposed, a branch of the modern #1 Family.

Interestingly however, it appears that based on the sequences of our study that the #25 Family (that tracing to the Brimmer Mare) is in same haplotype as the A4 Family and therefore shares a common matrilineal source. The most parsimonious interpretation of that might be to suggest that the founder mare of the A4 Family, The Medley Mare No. 2, whose own dam was known in the ASB only as an anonymous daughter of the imported stallion Sterling (1791), is in fact a member of the #25 Family that merely “lost her papers” on her ancestress’ importation into America.

This finding touches on the underlying dogma in both scientific and lay thoroughbred pedigree circles that highly performing American thoroughbred families must trace their maternal lineage to mares that were included in the GSB, as to explain their apparent racing class. The sequences of the A4 and #25 Family show they are more closely related than previously thought (and more distantly maternally related to the #1 Family), but given the diversity of equine populations that founded the thoroughbred and the mutations differing between the families, it is more likely that the A4 and #25 Family share a more distant common ancestor that pre-dates the GSB and other record keeping.

The fact that over 70 individual mtDNA samples studied specifically point to a well conserved female family that traces by stud book record to a single mare is astounding. While there are clearly distant errors in the General Stud Book and other thoroughbred registries around the world, generally speaking the accuracy of most of the branches of this single family and mitochondrial haplotype we examined is a broader testament to all the individuals involved during the past 300 years as stewards of this breed. In examining the racing class and history of the samples in this study it is apparent that the descendants of the samples tracing to the #1 family were kept in a tremendous range of circumstances (one in particular that excelled over hurdles in the 1940s in Ireland for a few generations) and in a wide range of geographic locations, giving greater appreciation to the remarkable Stud Book accuracy of this matrilineal haplotype.

The limitations of any study on mitochondrial DNA is of course that it cannot discount the possibility that there are Stud Book anomalies caused by ‘substitutions’ of animals with the same haplotypes (i.e. two foals out of closely related mares), which are obviously not detectable, and also the possibility of the sire of a distant ancestor being incorrect. However as our study showed, it can begin to pinpoint where an error looks to have been made in the recording of the female offspring of an ancestor in a pedigree, and also identify mitochondrial mutations that are unique to a particular “family” making classification possible. In an era where the cost of sequencing full mitochondrial genomes is becoming more and more affordable, the sequencing and proper classification of all racehorses is possible, and something from which the entire industry would benefit. ■

#### TREGONWELL'S NATURAL BARB MARE ... Family 1

Place's White Turk Mare (f Place's White Turk)

Taffolet Barb Mare (f Taffolet Barb)

Byerley Turk Mare (f Byerley Turk)

Darley Arabian Mare (f Darley Arabian)

**BONNY LASS** (b f 1723 Bay Bolton) ... Family 1-a

Partner Mare (b f 1735 Partner)

Julia (b f 1754 Blank)

**PROMISE** (br f 1768 Snap) ... Family 1-d

| **PRUNELLA** (b f 1788 Highflyer) ... Family 1-e

| **PENELOPE** (b f 1798 Trumpator) ... Family 1-o

| | **WEB** (b f 1808 Waxy) ... Family 1-s

| | | Fillagree (ch f 1815 Soothsayer)

| | | | Sister To Cobweb (ch f 1820 Phantom)

| | | | Odessa (ch f 1833 Sultan)

| | | | Flax (b f 1855 Surplice)

| | | | **QUEEN BERTHA** (b f 1860 Kingston) ... Family 1-w

| | | | Cobweb (b f 1821 Phantom)

| | | | Clementina (b f 1844 Venison)

| | | | Lady Blanche (b f 1856 Stockwell)

| | | | Lady Caroline (b f 1861 Orlando)

| | | | Casuistry (b f 1876 The Miner)

| | | | Inchbonny (b f 1883 Sterling)

| | | | Paradoxical (bl f 1891 Timothy)

| | | | Doxa (bbr f 1901 Melton)

| | | | Lady Of Pedigree (b f 1910 St. Denis)

| | | | Helene De Troie (b f 1916 Helicon)

| | | | **LA TROIENNE** (b f 1926 Teddy) ... Family 1-x

| | | | **TRAMPOLINE** (ch f 1825 Tramp) ... Family 1-t

| | | | Glencairne (b f 1838 Sultan)

| | | | Glengowrie (br f 1851 Touchstone)

| | | | **MAID OF THE GLEN** (b f 1858 Kingston) ... Family 1-u

| | | | Wire (br f 1811 Waxy)

| | | | Vinegar (bl f 1832 Picton)

| | | | Mustard (br f 1836 Philip The First)

| | | | Clarinda (bbl f 1846 Sir Hercules)

| | | | Nightingale (bl f 1857 Mountain Deer)

| | | | **HILARITY** (b f 1871 King Tom) ... Family 1-p

| | | | Waltz (ch f 1822 Election)

| | | | Morisca (b f 1826 Morisco)

| | | | Zillah (ch f 1835 Reveller)

| | | | **THE PRAIRIE BIRD** (b f 1844 Touchstone) ... Family 1-r

| | | | **PAWN** (b f 1808 Trumpator) ... Family 1-f

| | | | Pawn Junior (br f 1817 Waxy)

| | | | Delhi (bl f 1838 Plenipotentiary)

| | | | **ELLEN HORNE** (br f 1844 Redshank) ... Family 1-j

| | | | Paradigm (br f 1852 Paragone)

| | | | **PARAFFIN** (b f 1870 Blair Athol) ... Family 1-l

| | | | **FOOTLIGHT** (b f 1876 Cremorne) ... Family 1-m

| | | | Illuminata (br f 1877 Rosicrucian)

| | | | **CHELANDRY** (b f 1894 Goldfinch) ... Family 1-n

| | | | **ROUGE ROSE** (ch f 1865 Thormanby) ... Family 1-k

| | | | **PROBLEM** (ch f 1823 Merlin) ... Family 1-g

| | | | Io (ch f 1836 Taurus)

| | | | Sunflower (b f 1847 Bay Middleton)

| | | | Sunbeam (b f 1855 Chanticleer)

| | | | **SUNSHINE** (b f 1867 Thormanby) ... Family 1-h

| | | | **SUNRAY** (bbr f 1874 King Of The Forest) ... Family 1-i

Princess (b f 1769 King Herod)

Puzzle (b f 1778 Matchem)

Hornby Lass (b f 1796 Buzzard)

**MOREL** (ch f 1805 Sorcerer) ... Family 1-b

**MUSTARD** (ch f 1824 Merlin) ... Family 1-c