Speed Gene Background Essay

Introduction

Thoroughbred racehorses race in different distance categories ranging from 1,000 m (0.6 mile) to 4,000 m (5.5 miles) and, like human sprinters and marathon runners, are rarely capable of racing at a high level in multiple distance categories. Usain Bolt, one of the greatest modern day 100 m sprint athletes, would not perform well in long distance races against the top marathon runners and vice versa. It's the same for racehorses. Racehorses are generally considered to either be 'sprinters' (1,000 – 1,400 m), 'middle-distance' (1,600-2,000 m), or 'stayers' (>2,000 m).

Generally racehorses are either sprinters or stayers and, rarely, can they be champions in both types of races. Attempts to win at the most competitive level across the range of distances had not been successful for about 40 years until *American Pharaoh* won "The Triple Crown" race series in 2015. The Triple Crown is a series of races for three-year-old horses that test their ability to win the highest caliber races in the region. In the USA, the Triple Crown races are made up of the Preakness Stakes (1,900 m), the Kentucky Derby (2,000 m), and the Belmont Stakes (2,400 m).

What is the 'Speed Gene'

In 2010 a team of Irish scientists discovered the 'Speed Gene' and developed a genetic test (Speed Gene Test) which is now used by horse owners, breeders, and trainers around the world to predict whether a horse will have a better chance of winning races as a sprinter or as a stayer. The Speed Gene is the name given to the *myostatin* (*MSTN*) gene in horses. In other animal species like cattle, sheep, pigs, mice, and dogs, mutations in the *MSTN* gene cause unusual muscle characteristics.

Myostatin (*MSTN*) the 'Speed Gene'

"Double muscled" Belgian Blue bull
- Homozygous for a mutation in *MSTN* gene

European Champion 'sprint' racehorse ***Dawn Approach***
- Homozygous for a mutation in *MSTN* gene (C/C genotype)

'Bully' whippet dog
- Homozygous for a mutation in *MSTN* gene

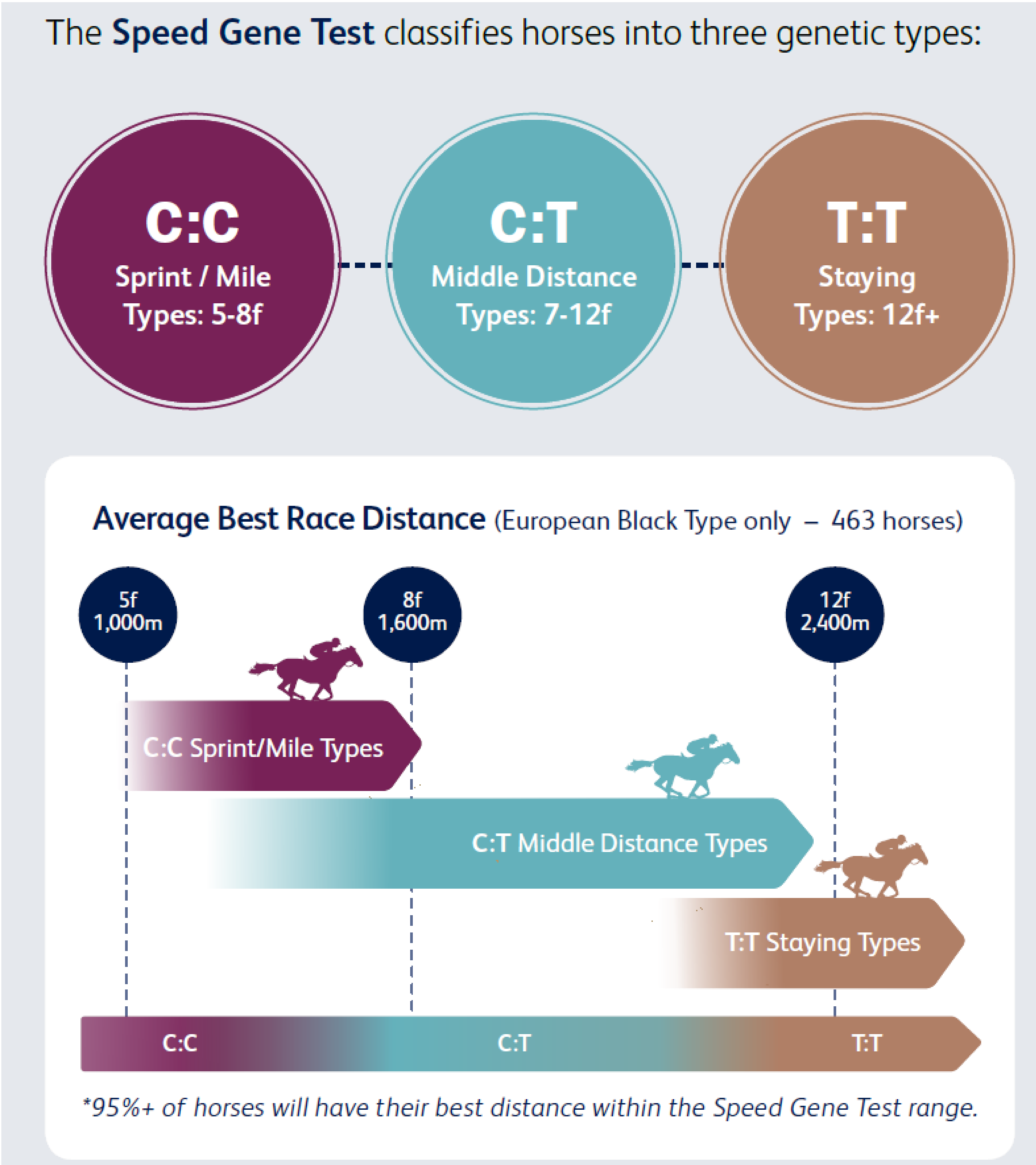
The 'Mighty mouse' (left)
- Gene edited 'knockout' mouse has no functional *MSTN* gene

In the Thoroughbred racehorse a mutation in the *MSTN* gene has a major effect on:

- Muscle hypertrophy - Skeletal muscle growth
- Muscle fiber type differentiation - The proportion of fast twitch (anaerobic glycolytic) muscle fibers required for short bursts of power and the proportion of slow twitch (aerobic oxidative) muscle fiber types required for stamina

Examining differences in the DNA of the *MSTN* gene allows an accurate prediction of a horse's best racing distance, based on the combination of "C" and "T" alleles at a particular position in the gene. The C-allele is linked to (a marker for) the causative mutation that results in greater muscle mass, more fast twitch fibers and ultimately greater speed. The 'T' allele is the 'wild-type' (i.e. no mutation) that is found in most other horse breeds (except the Quarter Horse). Since each Thoroughbred horse has two copies of the gene, - one inherited from each parent - there are three possible combinations of the alleles, known as genotypes. The three possible genotypes are C/C, C/T and T/T. The genotype depends on which combination of the two alleles has been inherited from each of the parents. The combination of alleles occurs at fertilization of the egg and sperm cells. Scientists have shown that the genotype can predict the race distance type.

- C/C genotype (homozygous for the C-allele) → Sprint/Mile Types
- C/T genotype (heterozygous, one C-allele and one T-allele) → Middle Distance Types
- T/T genotype (homozygous for the T-allele) → Staying Types



The ability to perform well at certain race distances is almost entirely determined by the genotype of the *MSTN* gene, but there are other genes that also have a small moderating effect on their best distance. It is very unusual for a single gene to have such a large effect on a 'complex' trait but the scientific evidence points to a major influence of *MSTN* on sprint racing.

The statistical association of the Speed Gene with race distance is as strong as the relationship between the androgen receptor gene and the chance of a man going bald, one of the strongest known links in human genetics.

Breeding

Traditionally, pedigree records (the ancestors of the horse) and conformation characteristics (what the horse looks like) have been used by Thoroughbred horse breeders and trainers to try to identify the best race distance for an individual horse at an early age or to identify the most suitable breeding choices to produce either sprinters or stayers. Race distance ability can also be assessed by the trainer and jockey during trials once the horse is in active training and by trial and error by having the horse race in different distance races.

Pedigree is not always a good indicator of distance because a C/T mare crossed with a C/T stallion can produce both sprinters and stayers. Because of this, breeders and trainers don't always get it right. So genetic testing for the Speed Gene can take away a lot of the guesswork.

Possible mating outcomes for each Speed Gene type:

Dam \ Sire	C:C	C:T	T:T
C:C	100% C:C	50% C:C 50% C:T	100% C:T
C:T	50% C:C 50% C:T	25% C:C 50% C:T 25% T:T	50% C:T 50% T:T
T:T	100% C:T	50% C:T 50% T:T	100% T:T

Selection for speed

There has been recent intense selection for speed in some horse breeds. Scientists looked at the genetics of 36 horse breeds from around the world, and found that the *MSTN* gene was under strong selection in the Quarter Horse breed. The Quarter Horse is an American horse breed that has been bred to be very fast in ¼ mile races (very short sprints). They also excel in rodeo sports like reining, cutting, barrel racing, calf roping, and other western riding events, that require agility and short bursts of speed. Most Quarter Horses have the C/C (sprint) genotype.

In 2012, scientists took DNA from the teeth and bones of some historically important Thoroughbreds that were born between 1877–1898. Their skeletons had been kept in museums because they were the greatest racehorses of the time. The racehorse *Eclipse* was one of the champion racehorses of the time – because he dominated on the racecourse there is a phrase that is still used today "*Eclipse first and the rest nowhere.*" The DNA from *Eclipse*'s bones tells us that he had the T/T genotype. The *MSTN* genotype distribution was very different historically than what it is now in the current population of racehorses. The champion racehorses at that time (late 1800s) all had T/T genotypes. This is because most racing in the nineteenth century was over much longer distances of 4-miles and this required horses to have a lot of stamina (rather than speed). Now racegoers prefer fast, sprint races and there has been a change in the distances of races in the last 100-150 years.

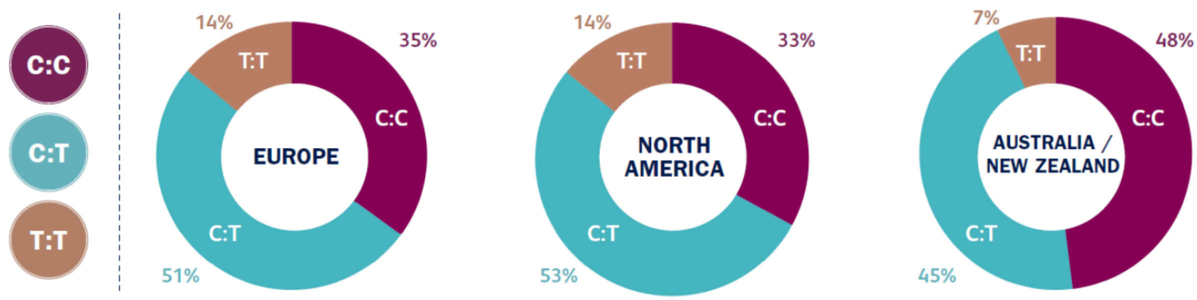
Dawn Approach, who features in *EQUUS: Chasing the Wind*, has the C/C genotype. He was the undefeated European Champion racehorse in 2012. He won the English 2,000 Guineas by five lengths and is now a stallion producing champion foals. His father *New Approach*, who was World Champion in 2008, has the C/T genotype.

New Approach can produce C/C, C/T and T/T foals because he has the C/T genotype. But *Dawn Approach* will only be able to produce C/C and C/T foals because he does not have a T-allele to pass on to his foals.

Allele frequencies in different regions
Without knowing the genotype of their horses, skilled horsemen have been unknowingly selecting for the Speed Gene for centuries as short distance races have become more popular. The C-allele is now at much higher frequency in modern day racehorses than it was 100-150 years ago (when the champions were all T/T).

Nowadays, the high frequency of the C-allele is even more evident in some regions of the world where there is very strong selection for shorter distance racing. In Australia, where the highest prize money is for sprint races, only 7% of racehorses now have the T/T genotype, but almost 50% of racehorses have the C/C genotype.

‘Speed gene’ genotype distribution in different racing regions of the world. The highest proportion of C/Cs are in Australia where there is more prize money for sprinters.



What does genetic selection do?
Selection in a population for a favorable / advantageous trait can very quickly alter the genotype distribution for the gene that is contributing to the trait.

Hardy-Weinberg Equilibrium
The Hardy-Weinberg equilibrium describes how allele frequency in a population will remain constant in the absence of genetic selection and other disrupting factors. In order to maintain this equilibrium the following conditions must apply:

- Mating must be random. There cannot be sexual selection for preferred traits.
- There are no mutations introducing new alleles into the population.
- The frequencies of alleles that positively or negatively impact organisms’ reproductive success isn’t impacted by natural selection.
- There is no gene flow. Organisms do not immigrate into or emigrate from the population.
- The population is large. In small populations allele frequencies grow higher or lower by chance. This is known as genetic drift.

These conditions rarely occur in nature. Therefore, the Hardy-Weinberg equilibrium describes an idealized state. Since genetic variations do occur naturally they can be measured as changes from the equilibrium.

Hardy-Weinberg Equation

The Hardy-Weinberg equilibrium can be expressed as an equation. If examining a gene which has two alleles, “A” and “a” the equation is as follows:

$$p^2+2pq+q^2 = 1$$

p = the frequency of the "A" allele
q = the frequency of the "a" allele

Therefore:

p^2 = the frequency of the homozygous genotype AA
2pq = the frequency of the heterozygous genotype Aa
q^2 = the frequency of the homozygous genotype aa

For example, the allele for hitchhiker's thumb, a thumb that can bend backwards, is recessive compared to straight thumbs. In a population the frequency of homozygous straight thumbs is 36 percent. What percentage of the population has straight thumbs?

$$p^2 = .36$$
$$p = \sqrt[2]{p2}$$
$$p = \sqrt[2]{.36}$$
$$p = .6$$
$$p + q = 1 \text{ so } q = .4.$$

We know p^2 = .36 so we need to calculate 2pq.

$$2pq = 2 X .6 X .4 = .48$$
$$.36 + .48 = .84$$
$$.84 X 100\% = 84\%$$

In “The Genetics and Evolution of Speed” handout, you will use the Hardy-Weinberg equation to further explore how the frequency of the Speed Gene has changed over time.

Acknowledgments
This essay was generously provided by Prof Emmeline Hill.
Professor of Equine Genomics, University College Dublin, Ireland
Chief Science Officer, Plusvital Ltd., Dublin, Ireland

Scientific references

Hill et al 2010 - A sequence polymorphism in MSTN predicts sprinting ability and racing stamina in thoroughbred **horses.** **Hill EW**, Gu J, Eivers SS, Fonseca RG, McGivney BA, Govindarajan P, Orr N, Katz LM, MacHugh DE. PLoS One. 2010 Jan 20;5(1):e8645. doi: 10.1371/journal.pone.0008645. Erratum in: PLoS One. 2010;5(1). doi: 10.1371/annotation/de9e11b9-eb92-4ee5-a56a-908e06d1ed6c. MacHugh, David [corrected to MacHugh, David E]. PMID: 20098749 **Free PMC Article**

Bower et al 2012 - The genetic origin and history of speed in the Thoroughbred racehorse. Bower MA, McGivney BA, Campana MG, Gu J, Andersson LS, Barrett E, Davis CR, Mikko S, Stock F, Voronkova V, Bradley DG, Fahey AG, Lindgren G, MacHugh DE, Sulimova G, **Hill EW**. Nat Commun. 2012 Jan 24;3:643. doi: 10.1038/ncomms1644. PMID: 22273681

Petersen et al 2013 - Genome-wide analysis reveals selection for important traits in domestic **horse** breeds. Petersen JL, Mickelson JR, Rendahl AK, Valberg SJ, Andersson LS, Axelsson J, Bailey E, Bannasch D, Binns MM, Borges AS, Brama P, da Câmara Machado A, Capomaccio S, Cappelli K, Cothran EG, Distl O, Fox-Clipsham L, Graves KT, Guérin G, Haase B, Hasegawa T, Hemmann K, **Hill EW**, Leeb T, Lindgren G, Lohi H, Lopes MS, McGivney BA, Mikko S, Orr N, Penedo MC, Piercy RJ, Raekallio M, Rieder S, Røed KH, Swinburne J, Tozaki T, Vaudin M, Wade CM, McCue ME. PLoS Genet. 2013;9(1):e1003211. doi: 10.1371/journal.pgen.1003211. Epub 2013 Jan 17. PMID: 23349635 **Free PMC Article**

Rooney et al 2018 - The "speed gene" effect of myostatin arises in Thoroughbred **horses** due to a promoter proximal SINE insertion. Rooney MF, **Hill** EW, Kelly VP, Porter RK. PLoS One. 2018 Oct 31;13(10):e0205664. doi: 10.1371/journal.pone.0205664. eCollection 2018. PMID: 30379863 **Free PMC Article**

Hill et al 2019 - The contribution of myostatin (MSTN) and additional modifying genetic loci to race distance aptitude in Thoroughbred **horses** racing in different geographic regions. **Hill EW**, McGivney BA, Rooney MF, Katz LM, Parnell A, MacHugh DE. Equine Vet J. 2019 Jan 3. doi: 10.1111/evj.13058. [Epub ahead of print] PMID: 30604488

