INTRODUCTION

Wheat is a *rabi* cereal crop for world because among staple crops, its grains are primarily consumed by humans. Wheat is originated from the eastern Mediterranean or the Middle East in Eurasia. Wheat belongs to the grass family, *Poaceae*. It is one of the most ancient domesticated crops extensively cultivated and consumed, covering more than 200 million hectares of land all over the temperate, Mediterranean and sub-tropical region of both Northern and Southern hemisphere (http://faostat.fao.org/2012). *T. aestivum* accounts for almost 95% of the wheat grown in the world, with 5% being durum wheat (*T. durum*) (Peng et al., 2011). Hexaploid wheat has the largest genome of approximately 17 billion nucleotides, 25%-30% gene duplication and present calculation of gene range from 77000 to 295,900 (Moolhuijzen et al., 2007). Moreover, wheat genome has 75-90% repetitive DNA makes its genome assembly more complicated. Bread wheat is the cheapest source of carbohydrates and proteins, used as food by one third population of the world. Wheat is one of the major food crops both in terms of area wise and production. The wide spread cultivation of the crop across the globe is largely due to high versatility of its genome. The variations in A, B and D genomes of wheat enable it to adapt wide agro-climatic conditions. Although India has a track record of producing more than 90 million tones continuously for a period of three years and reaching its historic peak (95.91 mt) in the year 2013. World’s population is estimated to reach 9.6 billion by 2050 and wheat production will have to play crucial role in food security and the global economy. The world Bank estimates that global wheat production must increase by 60 percent to meet rising demand (Tilman et al., 2011; FAO, 2009; OECD/FAO, 2012). The seed number and seed weight are critical yield components of wheat. Starch accounts for 65-75% of wheat seed weight, is a key determining factor of wheat yield. Starch is known to be an important carbohydrate and the primary energy source for plants, having various industrial applications (Slattery et
In non-photosynthetic starch-storing plant organs, starch is synthesized from the products of photosynthesis imported into the developing storage organ from the leaves. In cytosol and mitochondria sucrose catabolism produces the substrate for starch synthesis, glucose 6-phosphate and ATP. Specific translocators in the plastid envelope, namely the glucose 6-phosphate/phosphate transporter and the ATP/ADP transporter, transport it into plastid from cytosol (Neuhaus and Wagner 2000). Inside plastid, glucose 6-phosphate is converted to glucose 1-phosphate by phosphoglucomutase and then ADP glucose is made from glucose 1-phosphate and ATP by ADP glucose pyrophosphorylase (AGPase). For starch synthesis the glucose 6-phosphate/phosphate transporter may also be capable of transporting glucose 1-phosphate into plastids. ADP glucose in the plastid is used by various different isoforms of starch synthase to elongate the starch polymers amylose and amylopectin. These polymers are then branched by multiple isoforms of starch branching enzyme. The synthesis of starch granules also requires other enzymes including isoamylase. The synthesis of starch in chloroplast is analogous to that in non-photosynthetic organs except the glucose 6-phosphate and ATPs are made photosynthetically inside the plastid. The plants AGPase are regulated through allosteric activation by 3-PGA (Phospho glyceral acid) and inhibition by orthophosphate. Any upward modulation of AGPase activity will have direct impact on the net starch synthesis. AGPase consists of two large subunits and two small subunits, each of which is encoded by distinct genes. The enzyme is now known to be largely extra-plastidial (85-95% cytosolic) in cereal endosperm, conversely plastidial in other cereal tissues and in all tissues of non-cereal plants. Recent studies suggested that distinct cytosolic and plastidial forms of AGPase that are encoded by separate large- and small- subunit genes exist in all cereal endosperms.
However, the availability of the Expressed Sequence Tag (EST) resource has greatly advanced genomic research in wheat. ESTs have been demonstrated to be valuable resource for genome analysis, identification of candidate genes, gene annotation and comparative genomics (Ergen and Budak, 2009). In plant genetic research, marker-assisted selection (MAS) offers a tactic for accelerating the process of wheat breeding. The genetic maps constructed using markers like RFLPs, SSRs and DArT markers are most low density maps, thus making them inappropriate for high resolution quantitative trait loci (QTL) analysis or association mapping. Therefore, there is acute need of a marker system which will allow development of high-density molecular maps for complete association genetic studies (Gupta et al., 2008). Single nucleotide polymorphisms (SNPs) are used to fulfill the present requirement for marker system. Using bioinformatics approach, EST based SNP have been mined in large number of crops EST databases including barley (Duran et al., 2009), maize (Batley et al., 2003), rice (Feltus et al., 2004), and wheat (Somers et al., 2003).

Swift progress in genomics and bioinformatics in recent years has generated huge amount of wheat data. Some efficient and suitable algorithms/programs are essential to organize and access the desired data. Such programs are called as database management system (DBMS) (Silberschatz et al., 1998). Data stored in the database such as enzyme database enable the wheat researchers, all around the world to use this publically available information in the wheat breeding program to identify variety of interest like low or high yield. It will provide access to AGPase information in wheat and other crops.

An attempt was made by different researchers to get crystal structure of AGPase which allow its structural characterization but was unsuccessful due to difficulty of obtaining the enzyme in pure form. This confines the complete understanding of structure–function relationships of this enzyme. An effective approach to solve this snag is the use
of computational method to predict protein structure (Rost and Sander, 1993). Computational approaches act as time and cost efficient aid to experimental procedures. It narrows down the search space for better annotation and recognition of biological function and characteristics.

Notable innovations in platforms for omics-based research and application development provide vital resources to stimulate research in model and applied plant species. A combinatorial approach is now an effective tactic for clarifying molecular systems integral and increasing plant productivity. Bioinformatics platform and their associated databases are essential for effective design for making the best use of genomic resources. Due to complexity of wheat genome, information on structural and functional analysis of wheat AGPase is not much explored.

Therefore, considering the above perspective in view, the present study is proposed with the following objectives:

- Comparative gene analysis of wheat AGPase with other well studied AGPase using various bioinformatics tools
- *In silico* study of interaction between large and small subunit of wheat AGPase enzyme
- To explore SNPs in AGPase and their validation in selected wheat genotype
- Knowledge base bioinformatics database development of AGPase