The authors have examined a diverse set of 130 spring wheat genotypes for genotypic variation of spike-ethylene, spike dry weight and yield under heat stress (HT) and heat stress with silver nitrate application (HSN). This was done in the frame of a GWAS analysis to identify genetic regions underlying this variability.

The present study reports, indeed, on a field of increasing importance in the global wheat production. Formerly a serious problem particularly in Mediterranean environments, heat stress has nowadays become a problem also in temperate seasonal climates as a consequence of global warming, leading to yield reductions of up to 40%. Therefore, I fully agree with the statement of Reviewer 2 that the study reports on a field of great importance for wheat breeding.

The presented study represents a hypothesis-driven research approach. The inhibition of ethylene perception as a manner of lowering stress impacts on grain development in heat susceptible cultivars has been proven, and was accordingly reported in the present study. However, knowledge on the genetic regions underlying heat stress related processes, particularly concerning ethylene effects, in wheat spikes was, as far as I know, hitherto not available. Also in this point I agree with Reviewer 2. Finally, the revisions given by Reviewer 2 have been implemented into the revised manuscript in a satisfactory manner.

With regard to the overall quality of the work, the experimental design is conclusive (e.g. four field experiments and an additionally greenhouse study) and the data are well-analyzed. In this context, the first comment given by the Reviewer 1 criticizes the lack of a control experiment with plant grown under non-heat-stress conditions. In the first instance, I can understand this concern. However, I agree with author’s response that the concept of field evaluations for heat stress responses does not allow testing for non-heat control conditions, without implementing another growing season or geographic location. This, indeed, would introduce an experimental variation which may be even more critical than the used HSN condition, since the traits assayed (as most agronomical traits) show a relatively large genotype-by-environment interaction. A non-heat-treatment is feasible under greenhouse conditions, but such data must stay isolated since they cannot be related to the field data. The second concern of Reviewer 1 with regard to the spike sampling and ethylene measurement
has been answered satisfactory by implementing the necessary information in respective section of Materials and Methods.

Personally I have just some minor comments or suggestions.

Line 159: Experimental design and details. Here I would skip ... and details.

Line 180: Here the authors wrote "... during the vegetative period in 2013, 2014, and 2015". To my understanding the vegetative period is referred to the growth phase between germination and flowering. Maybe at this point the authors mean the respective vegetation or growing periods in 2013, 2014, and 2015. Finally the term vegetative period is used far too often, e.g. not necessary within the brackets in line 182.

In the same section it is written: …the average temperatures of the crop canopy … The term crop canopy in this context, is somehow confusing, at least for me, since canopy refers to the aboveground portion of a crop, often used in the context of forest ecology as tree crowns. Which growth period is exactly meant? Does this term refer to the flowering or grain development stages? Generally, I suggest that developmental stages or periods might be defined by an indicator such as the Zadoks growth scale.

In contrast to several other studies which investigate short-term heat stress (around three days of heat treatment), a phenomenon nowadays typical for temperate seasonal climates, in the present study tested for a long-term heat stress. It might be helpful to make this clear early in the text. There might be differences concerning the manner in which heat stress causes yield reduction. I remember that short heat stress impairs flower development and/or grain filling, while sustained heat stress (starting during vegetative growth) significantly reduces crop yield by accelerating plant senescence.

The authors report that a major SNP (kukri427 c15603-1116) at 69cM in chromosome 3B explained 10.5% of total spike-ethylene variation under field conditions. In this context, a very recent publication reports on a QTL on wheat chromosome 3B that affects the stability of grain weight in plants exposed to heat stress early in grain filling (Shirdelmoghlanloo et al., 2016). It may be interesting to check whether a genetic overlap exists.