CASE FOR SUPPORT: **Genomics-led improvement of biotic and abiotic stress tolerance in mustard rape for economic and environmental sustainability**

a) Research track record, specific expertise and infrastructure

University of York (UoY) Department of Biology was ranked 1st for impact in the latest Research Excellence assessment. It has an extensive infrastructure, including a Technology Facility and servers tailored to support the analysis of NGS data. **Ian Bancroft** holds the CNAP Chair of Plant Genomics. His group elucidated mechanisms by which *Brassica* genomes evolved and played a key role in the consortium that sequenced the genome of *Brassica rapa* (BrGSPC *Nature Genetics* 43:1035-1039, 2011). Recently, his group has focused on applying genomics technology to trait genetics, using Illumina sequencing to bringing about a step change in genome analysis (Bancroft et al *Nature Biotechnology* 29:762-6, 2011). Prof. Bancroft has an extensive record of successful coordination of collaborative projects (e.g. 2014 BBSRC BB/L011751/1 Broadening the genetic diversity underpinning seed quality and yield related traits in mustard rape and oilseed rape; 2014 BBSRC BB/L002124/1 BBSRC Renewable Industrial Products from Rapeseed (RIPR) Programme). **Andrea L. Harper** developed the technology termed Associative Transcriptomics (Harper et al *Nature Biotechnology* 30:798-802, 2012), which enables the identification of causative genes.

University of Delhi South Campus (UDSC) is one India's top 10 universities. It has strong teaching and research profiles and a broad range of facilities to support crop-oriented research. **Akshay Kumar Pradhan** has been intensely involved in improving mustard rape through introgression and transgenic technologies. A major accomplishment is the successful delivery of two hybrids in Indian Mustard (DMH1 and DMH11), which provide 15-30% increase in yield over the best national and regional check varieties. Two loci for white rust in *B. juncea* lines have been identified using bi-parental mapping populations. These rust resistance loci are being introgressed into popular Indian varieties. **Jagreet Kaur** works towards understanding the mechanisms underlying *Brassica juncea* resistance/susceptibility to *Alternaria brassicae* and *Sclerotinia sclerotiorum*.

The Punjab Agricultural University (PAU), the top Agricultural University in India, played a key role in increasing food grain production in Punjab, ushering in an era of Green Revolution in India. **S.S. Banga** has 3 decades of experience in pre-breeding and cytogenetics in *Brassica* oilseeds. He holds ICAR National Professor Chair and leads DBT Centre of Excellence on Brassicas. **Mehak Gupta** is Asst. Prof. (Cytogenetics) and recipient of ICAR best thesis award-2016. **Rimaljeet Nagra** is Asst. Prof. (DST INSPIRE Faculty) and recipient of ICAR best thesis award-2015. **Chhaya Atri** is Asst. Prof. with experience in wide hybridization and molecular genetics and was recipient of Endeavour Fellowship of Australia-2008. **Virender Sardana** is Professor of Agronomy, experienced in input use efficiency and abiotic stress resistance. **Sarwan Kumar** is Asst. Entomologist, experienced in screening against mustard aphid and cabbage caterpillar. **Sanjula Sharma** is a biochemist with experience in developing NIR equations and UPLC based analysis.

Directorate of Rapeseed-Mustard Research (DRMR) The Centre has responsibility to plan, coordinate and execute the research programmes through wide network of 22 main and sub-centres across the country, to augment the production and productivity of rapeseed-mustard. **P.K. Rai** is Director and an expert in white rust. **Pankaj Sharma** is an expert in stem rot. **P.D. Meena** is an expert in white rust and black spot. **V.V. Singh** is an expert in heat stress.

National Institute of Plant Genome Research (NIPGR) is a premier Institution for plant genomic research. It has extensive infrastructure to support technologies for sustainable agriculture. **Naveen C. Bisht** has expertise in *B. juncea* transgenics and manipulating glucosinolate pathway genes for enhanced nutritional value and disease resistance (Augustine et al. Plant Biotech J 11: 855-66, 2015; Kumar et al. PC&E 40: 2109-20: 2017).

ICAR-National Research Centre on Plant Biotechnology (NRCPB) is the leading national centre in the area of genomics and crop biotechnology. It participated in several international genome sequencing consortia and has extensive genomics facilities including an NGS data centre. **Ramcharan Bhattacharya** has extensive work on developing resistance to aphid and *Alternaria* in *B. juncea* (Koramutla et al *Scientific Rep*orts 6: 25883, 2016; Bhattacharya et al *Plant Science* 207: 88-97, 2013). **Mahesh Rao (MR)** developed alien introgression lines in *B. juncea* (Vasupalli et al *Molecular Breeding* 37:110, 2017). **Kishor Gaikwad** applies NGS in plant transcriptomics (Zhang et al *Plant Cell* 29:1952-69, 2017; *Nature* 485:635–641, 2012). **Navin Gupta** has expertise in

Sclerotinia resistance. **Anshul Watts** has expertise in broomrape (*Orobanche*) and a haploid inducer line (Watts et al. 2017 Front Plant Sci https://doi.org/10.3389/fpls.2016.02019).

Y.S. Parmar University (YSPU) is fully equipped for handling and maintaining virus isolates and for serological and molecular detection of plant viruses, viroids and phytoplasmas. **Anil K. Handa** has extensive experience of virus research and ELISA-based serological detection of viruses infecting plants (e.g. Bohra et al, *J. Mycol. Plant Pathol.*46: 273-286, 2017).

ICAR-Indian Agricultural Research Institute (IARI) is the premier agriculture research institute. It has ranked second by the UGC for post graduate teaching and advanced research in agriculture. **Lakshman Prasad** has extensive screening experience and amassed diverse collections of *Alternaria*, *Sclerotinia* (Choudhary and Prasad, *VEGETOS*, 25: 178-183, 2012) and white rust.

ICAR-National Bureau of Plant Genetic Resources (NBPGR), has a large collection of crop germplasm and manages the sustainable use of such economically important plant genetic resources. **Rajesh Kumar** has expertise in understanding genomic basis of diversity in traits related to seed composition and environmental stress resistance/tolerance in oilseed species.

Indian Institute of Technology Kharagpur (IITK) is India's highest-ranked institution for nurturing graduate employability. **Mrinal Kumar Maiti** has expertise in plant molecular biology & transgenics.

Central Arid Zone Research Institute (CAZRI) has the mandate to undertake basic and applied research on sustainable farming systems in the arid ecosystem. **Hans Raj Mehla** has expertise in high temperature tolerance.

University of Warwick (UoW) School of Life Sciences has facilities for plant research ranging from the molecular to field scale. **John Walsh** has considerable experience of brassica virus research, and is recognized as a world leader in this field. **Guy Barker** has expertise in CRISPR-Cas9 technology in Brassicas. **Graham Teakle** has expertise in viruses.

Rothamsted Research (RRes) are equipped for metabolomics enabling (*inter alia*) lipidomics. **Fred Beaudoin** has over 20 years' experience in metabolic engineering of plant lipid content and composition, (e.g. Ruiz-Lopez et al *Plant Biotechnology Journal* 15:837-849, 2017). **Jon West** is expert in multi-disciplinary crop protection research and a pioneer of Aerobiological sampling (e.g. Rogers et al *Plant Pathology* 58: 324-331, 2009). **Samantha Cook** has extensive expertise in behavioral ecology of insect pests, particular in the effects of plant waxes on feeding behavior.

The University of Leicester (UoL) Department of Genetics and Genome Biology is one of the top research Departments in the UK for genetics research. **Pat Heslop-Harrison** is a world-leader in molecular cytogenetics and chromosomal studies to integrate genomic data for the exploitation of and recombination in polyploid crop breeding (Ali et al *Heredity* 117, 114–123, 2016)

University of Essex (UoE) School of Biological Sciences at Essex houses a wide range of facilities for the growth and physiological measurement of plants. **Ulrike Bechtold** has expertise in network inference using gene expression to identify regulatory genes (Bechtold et al *Plant Cell* 28:345–366, 2016). **Phil Mullineaux** has a track record on the response of plants to environmental stress and associated signaling of drought responses (Exposito-Rodriguez et al Nat. Comm. 8:49, 2017). **Tracy Lawson** has extensive experience investigating stomatal physiology and photosynthesis in controlled and field environments (e.g. Vialet-Chabrand et al *Plant Physiol.* 174(2): 603-613, 2017).

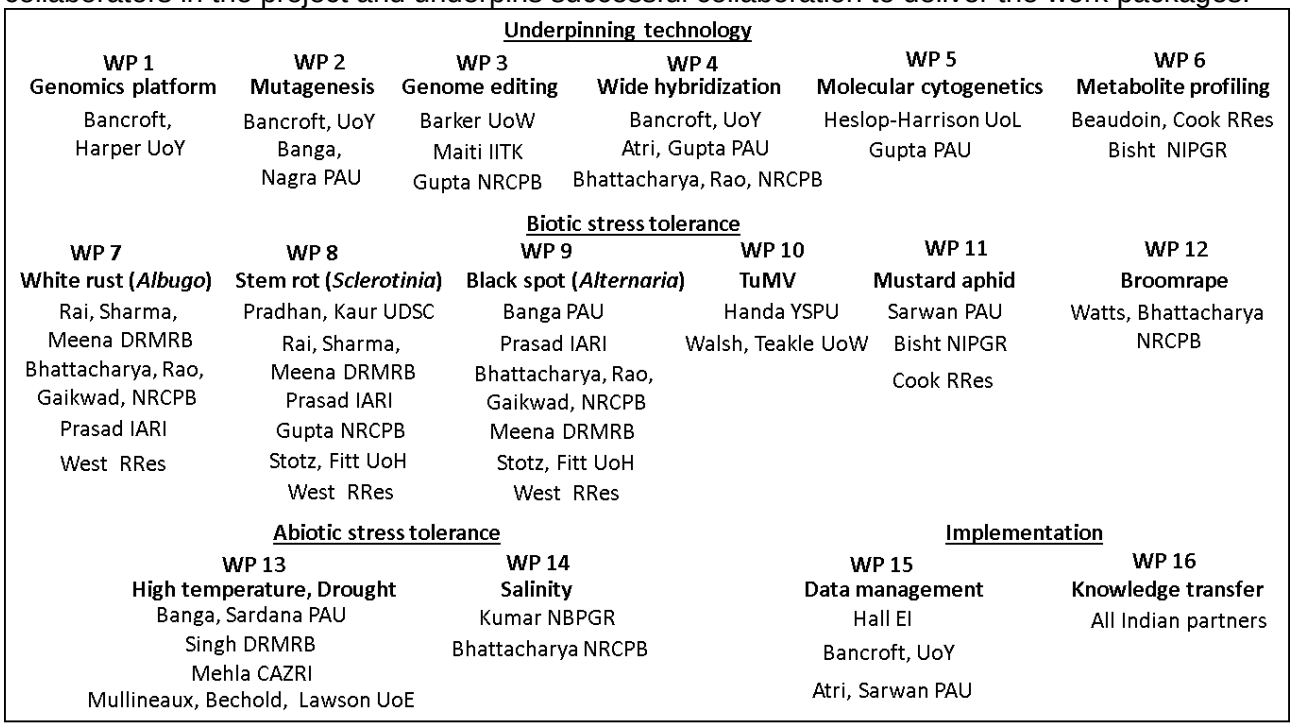
University of Hertfordshire (UoH) hosts a the new £50M Science Building that supports biological research, including a laboratory for microbiology and plant pathology, and provides plant growth facilities. **Prof Bruce Fitt** and **Dr Henrik Stotz** have respective expertise in epidemiology of and functional genomics of resistance against *Alternaria* spp., *Sclerotinia sclerotiorum* and extracellular fungal pathogens of crops (Stotz et al *Trends in Plant Science* 19: 491-500, 2014).

Earlham Institute (EI) is focused on exploring living systems by applying computational science and biotechnology to answer ambitious biological questions and generate enabling resources. **Prof Anthony Hall** is a leader in crop genomics (e.g. Brenchley et al *Nature* 491:705–710, 2012).

b) Strength of the scientific collaboration

We will build upon existing collaborations, expanded to include the expertise required to address, in an integrative way, the target traits. At its core is the genomics and bioinformatics expertise of the Bancroft group, which was developed for *B. napus* and already proven to be transferable to *B. juncea*

via CGAT BB/L011751/1, which is a collaborative project with Prof. Pradhan of University of Delhi South Campus. Both Profs Bancroft and Pradhan, the UK and India coordinators, have extensive networks of scientific contacts within the respective countries. This has guided the choice of collaborators in the project and underpins successful collaboration to deliver the work packages:



c) The proposed research

Background

Mustard rape (*Brassica juncea*) is the major oil crop in India. It is cultivated on various scales as a rain fed crop, with estimated production for 2016-17 of 8.0 Mt from 6.8 Mha being worth to the farmer ~Rs 30,000 crore (~GBP 3.7 Bn), maintaining India's position as the third largest producer of oilseed/mustard rape worldwide. However, around 20 Mt of edible oils must still be imported (Kumar et al 2009). Mustard rape has the potential to largely fill this gap, but policy decisions, such as the crop not benefitting from Minimum Support Price subsidies as enjoyed by wheat and pulse crops, means that it is in particular need of assistance to improve its economic viability.

Mustard rape yields have stagnated, despite the release of improved elite cultivars (including F_1 hybrids), as the crop faces many production challenges, particularly from the abiotic stresses drought, high temperatures and salinity, the diseases white rust (*Albugo candida*), stem rot (*Sclerotinia sclerotiorum*) and black spot (*Alternaria brassicae*), the potyvirus Turnip Mosaic Virus (TuMV), pests such as mustard aphid (*Lipaphis erysimi*), large white butterfly (*Pieris brassicae*) and root parasitism by broomrape (*Orobanche* spp.). Plants have defence compounds, including glucosinolates and surfaces waxes. Research has focussed on single characteristics (traits). This is inefficient and misses understanding of trade-offs and synergies. Only an integrated analysis will fully support the predictive breeding of more resilient mustard rape varieties, providing the consistent yields that bolster farmer confidence in the crop and its economic viability, and incorporating considerations of sustainability in the context of smallholders in India.

Recent advances in genome sequencing have led to the emergence of draft genome sequences for the allotetraploid species *B. juncea* (A and B genomes) and its diploid progenitor species *B. rapa* (A genome) and *B. nigra* (B genome) (Yang et al 2016). The published draft genome sequence of *B. juncea* is particularly poorly organized relative to the *B. juncea* genetic map (He et al *submitted*), meaning that genetic and molecular analyses of traits based on this resource will be misleading. However, genomic and bioinformatic technologies developed by the Bancroft group, originally for *B. napus* (oilseed rape; A and C genomes) (Trick et al 2009; Bancroft et al 2011) permit improvement of the organisation of genome sequences and underpin the development of a "pan-genome", achieved by interpolating genes specific to the allotetraploid species into a framework comprising the progenitor genomes rigorously mapped to represent their organisation in the allotetraploid (He et al 2015). A preliminary pan-genome resource for *B. juncea* has been developed

by the Bancroft group with some success (He et al *submitted*, Harper et al *in preparation*), but there is much scope for improvement. High quality genomic platforms will enable use of powerful approaches such as Associative Transcriptomics (Harper et al 2012) to understand the genetic basis of trait variation and identify candidates for causative genes, which can then be validated using emerging technologies such as genome editing and genome re-sequenced mutation panels.

The proposed Newton-Bhabha development programme aims to transfer and optimize UK genomics expertise for improvement of both economic and environmental sustainability of mustard rape. The selected traits (tolerance to biotic and abiotic stresses) address the funders' aims: "to enhance yield potential, increase tolerance to climatic stresses or poor-quality soils, or counter pests or diseases". The approaches taken, founded on UK expertise in genomics and bioinformatics to exploit genetic variation and Indian expertise in the breeding and trialling of mustard rape, complies with the stated purpose of the initiative: "to use genomic approaches to accelerate varietal improvement". Consistent with the aims of the call, we will take a genomics-led approach. To underpin this, the *Brassica juncea* genomic platform currently used by the University of York and University of Delhi South Campus as part of their current Crop Genomics and Technologies (CGAT) project will be updated to incorporate emerging genome sequences from *B. juncea* and its progenitor species. The platform will be used to support the trait-focussed activities of the consortium, modelled on the University of York-led "BBSRC Renewable Industrial Products from rapeseed (RIPR) programme". We will undertake an Associative Transcriptomics analysis of an expanded *B. juncea* genetic diversity panel showing variation for the traits being studied as well as exploiting specific research material already developed (such as introgression material, resistant research lines, etc.).

Aims

We aim to apply to mustard rape cutting-edge genomic approaches first developed for oilseed rape in order to underpin improved productivity, sustainability and profitability. Specifically, we will:

1. Establish an updated Associative Transcriptomics platform for *B. juncea*.
2. Establish a genome re-sequenced *B. juncea* radiation panel platform.
3. Establish CRISPR-Cas9 technology in *B. juncea*.
4. Establish transcriptome re-sequenced resources for *B. juncea* introgression lines.
5. Establish a molecular cytogenetic toolkit to support trait introgression into *B. juncea*.
6. Establish a profile resource for key stress-relevant metabolites in *B. juncea*.
7. Understand the genetic and molecular basis of white rust tolerance.
8. Understand the genetic and molecular basis of stem rot tolerance.
9. Understand the genetic and molecular basis of black spot tolerance.
10. Understand the genetic and molecular basis of extreme virus resistance.
11. Understand the genetic and molecular basis of aphid and butterfly tolerance.
12. Understand the genetic and molecular basis of *Brassica*-broomrape interaction.
13. Understand the genetic and molecular basis of temperature and drought tolerance.
14. Understand the genetic and molecular basis of salinity tolerance.
15. Ensure effective data management to support pre-breeding.
16. Ensure effective knowledge transfer.
17. Ensure effective project management.

Methodology

The research will be conducted *via* a set of work packages (WPs), each addressing a specific aim. WPs 1 to 6 result in the establishment of technology platforms for mustard rape, WPs 7 to 14 aim to undertake focused research and draw upon the technology platforms to identify the genetic and molecular bases of stress trait tolerance, WPs 15 and 16 aim to empower the development and adoption of improved cultivars based on the research and WP17 addresses project management.

WP1. Establish an updated Associative Transcriptomics platform for *B. juncea*

Milestones: **M1.1** Month 3 - Organisation of latest A and B diploid genome sequences as represented in *B. juncea* confirmed using VHDH population; **M1.2** Month 6 - Genes specific to *B. juncea* interpolated into the *Brassica* AB framework to establish the *Brassica* AB pan-genome resource; **M1.3** Month 12 - Germplasm for ~300-accession *B. juncea* diversity panel collected and amplified; **M1.4** Month 15 - Associative Transcriptomics platform tested; **M1.5** Month 24 - Trait correlation analysis, training and Associative Transcriptomics using project trait data completed.

The innovative association genetics approach termed Associative Transcriptomics (AT), which combines GWAS based on transcriptome SNPs with the association of transcript abundance with trait variation has proved remarkably successful for identifying candidate causative genes in various species (e.g. Harper et al 2012; Miller et al 2016; Sollars et al 2017). Germplasm representing an expanded *B. juncea* diversity panel of ~ 300 accessions will be collected from partners UDSC, PAU, NBPGR, amplified in India and distributed amongst partners for trait analysis. In parallel, frozen leaves from plants being grown for amplification will be harvested simultaneously (as described in Harper et al, 2012) and shipped to UoY Bancroft for RNA extraction. mRNA sequencing will be contracted to provide ~20M 150 base PE reads for each of the ~300 accessions.

A pan AB genome resource will be developed by UoY Bancroft for *B. juncea* based on the genome sequences of the diploid progenitors (*B. rapa* and *B. nigra*) with interpolation of genes specific to *B. juncea*, as described previously for *B. napus* (He et al, 2015). The ordered CDS gene models from this will be used as reference sequences for scoring of transcriptome SNPs and quantification of gene expression across the diversity panel, as described previously for *B. napus* (Harper et al, 2012). Population stratification for the new panel will be calculated by UoY Harper from these SNPs and they will test the new *B. juncea* AT platform using the new functional genotypes and population data, along with existing trait data from a subset of the population as used in CGAT BB/L011751/1. Indian scientists visiting UoY will be trained in AT via "live" analysis of emerging project trait data from the new panel. Results will be collated and graphed for return to project partners for identification of candidate control genes by the trait experts. As the hub for trait data from the diversity panel, UoY Bancroft will develop and distribute to partners a comprehensive trait correlation matrix to aid identification of relevant relationships between traits.

WP2. Establish a genome re-sequenced *B. juncea* radiation panel platform

Milestones: **M2.1** Month 16 - Genome re-sequencing data available for ~400 lines; **M2.2** Month 18 - Variant scoring complete and panel available for use.

Ionising radiation has a long history of successful use for trait-led mutation breeding (https://mvd.iaea.org/#!Home). Building on early use of genome re-sequencing to investigate mutation spectra (Belfield EJ et al 2012), the Bancroft group has revealed that deletions, duplications and frame shift mutations are sufficiently abundant (i.e. 1% of all genes contain a frame-shift mutation at optimum radiation dosage) to make this cost-effective in allotetraploid *Brassica* species (Havlickova et al unpublished), underpinning genomic-led mutation breeding. As well as circumventing the need for locus-specific PCR as required for "TILLING", which can be difficult in polyploids, a much broader range of mutations are caused than with chemical mutagens such as EMS. A *B. juncea* radiation panel is already in development at PAU using a Canola-type accession. Leaves from ~400 of the optimal dose (1500 Gy) M_2 lines will be collected by PAU Banga, shipped to UoY Bancroft for genomic DNA extraction. DNA sequencing will be contracted to provide ~10x re-sequencing data. Variant detection will use BWA and custom scripts as developed for *B. napus*, using the AB pan-genome DNA genome models as reference sequences. A spreadsheet of variants detected across the panel, along with images showing larger deletions and duplications, will be distributed to partners and used by them for the selection of lines to test trait-control hypotheses.

WP3. Establish CRISPR-Cas9 technology in *B. juncea*

Milestones: **M3.1** Month 18 - CRISPR-Cas9 technology established in *B. juncea*.

At UoW a number of specialised plasmids have been developed for CRISPR cas9 use in both *B. rapa* and *B. oleracea*. These will be tested with *B. juncea* lines by Drs Maiti and Gupta. 100 seeds from each line will be grown to cotyledon stage and dipped in agrobacterium transformed with assistant and helper plasmids. Callus will be induced and somatic embryos grown on to plantlet stage. Verification of transformation events will be monitored using GUS fusions. Initial transformation will target the *BjFAE1* genes for knock-out as an exercise to provide training (in Warwick) to the Indian partners. The plants will be grown to maturity and the seeds tested for the phenotypic impact expected for successful genome editing: reduction in seed erucic acid content. Once validated, genome editing will be used, in India, to test candidate genes across the programme.

WP4. Establish transcriptome re-sequenced resources for *B. juncea* introgression lines

Milestones: **M4.1**Month 6 - Plant lines for re-sequencing identified; **M4.2** Month 24 - Transcriptome re-sequencing of alien introgression lines completed and introgressions delineated.

Alien introgression has enormous potential for crop improvement (Zamir 2001). By understanding precisely which genomic regions are introduced following wide crossing and their

molecular impacts, we can understand better the molecular bases of the traits and improve the efficiency of selection using molecular markers. Pilot investigation by the UoY Bancroft and PAU Banga groups (He et al *unpublished*) have shown that introgressions into *B. napus* and *B. juncea* are not silenced, allowing mRNAseq to be used to simultaneously delineate the introgressions and explore molecular impacts. A panel of introgression lines of relevance to the program will be identified. For the first ~300, leaf tissue will be frozen and shipped to UoY Bancroft for RNA extraction. mRNA sequencing will be contracted to provide ~20M 150 base PE reads for each sample. These data, plus those from further lines sequenced by the source groups, will be analyzed by UoY Bancroft to delineate introgression segments essentially as described for homoeologous exchanges (He et al 2016) but using the AB pan-genome CDS gene models plus a third colour channel for the introgressed genome. Transciptome assembly will be used to provide gene models for the alien introgression where that genome has not been sequenced. The materials are *B. juncea* introgression lines derived through wide hybridization with: *Diplotaxis erucoides* (NRCPB), screened for resistance to *Alternaria* and white rust; *Brassica fruticulosa*, *Erucastrum abyssinicum*, *Diplotaxis tennuisiliqua* (PAU), screened for resistance to diseases, drought and heat stress.

WP5. Establish a molecular cytogenetic toolkit to support trait introgression into *B. juncea*
Milestones: **M5.1** Month 12 - Development of gene-targeted *in situ* hybridization to localize physically genes of interest on identified chromosomes; **M5.2** Month 24 - Parameters for efficient transfer between genomes of small chromosome segments determined.

In India, the only route-to-market to exploit variation present in related species is through hybridizing (crossing and backcrossing) a combination of introgression characteristics from the donor species and subsequent recombination to reduce transfer of undesirable characters (linkage drag). Bioinformatics approaches on sequence data will be used to develop chromosome- and genome-specific markers to characterize introgression and translocation lines, and design probes for in situ hybridization to localize physically genes of interest and monitor their transmission. We will then exploit fertile and sub-fertile hybrid materials, use probes from above on meiotic material to identify causes of sub-fertility and identify physical sites of recombination using synaptonemal complex antibodies (cf Sepsi, Heslop-Harrison et al 2017) in tetraploids and hybrids. These studies will enhance approaches to efficient recombination and identify lines where chromosome pairing is satisfactory and small chromosome segments can be transferred, complementing marker-based tests. We will use variation between lines to maximize the number of different intercalary and intergenomic recombination events enabling exploitation of genetic variation without linkage drag.

WP6. Establish a profile resource for key stress-relevant metabolites in *B. juncea*
Milestones: **M6.1** Month 15 - *B. juncea* diversity panel analysed for compositional traits: leaf glucosinolates, leaf and stem wax, leaf sucrose and leaf reflectance; **M6.2** Month 27 - Candidate genes for control of compositional traits impacting pest/disease tolerance identified.

The glucosinolate–myriosinase system, also known as mustard oil bomb, constitutes the major defense arsenal of crucifers. The NIPGR component will investigate the molecular-genetic basis of glucosinolate variation across the ~300-accession *B. juncea* diversity panel, which will be grown under greenhouse and field conditions. Analysis of total and component glucosinolates (measured as desulfoglucosinolates) will be performed in young leaf (4-6 weeks) and mature seeds. AT of glucosinolate profiles will be undertaken (UoY Harper), which will identify the candidate genes controlling the variation of glucosinolate content and types, especially the species-specific aspects.

In addition to protecting plants against abiotic stresses, cuticular wax is involved in interactions with pest and pathogens. Variation in wax content, composition and structure is associated with resistance to aphids and to *Sclerotinia sclerotium* in *B. napus* (Eigenbrode et al., 2000; Ni et al., 2013). Leaf and stem wax content and composition will be analysed across the same *B. juncea* panel at RRes, using a reproducible, high throughput method developed as a part of the BBSRC RIPR sLoLa program. Plant material cultivated in India will be extracted with chloroform and dry wax samples will be shipped to RRes. The data generated will be submitted to UoY for AT analysis to identify genes controlling wax quantitative and qualitative traits variation.

Sucrose is a known feeding stimulant in many insect species including known pests of *Brassica* (Herve etl al 2014). Leaf sucrose content will be assayed across the diversity panel at RRes using young basal leaves harvested from 4-to-6-week-old plants using a sucrose assay kit (K-SUCGL, Megazyme Wicklow, Ireland) following the supplied instructions.

Leaf spectral reflectance can be affected by a number of physiological factors, attack by pest and pathogens, and physicochemical factor such as epicuticular wax. Differential Leaf spectral reflectance due to variation in cuticular wax is expected to affect plant attractiveness to pests that use visual cues (colour) for landing (Doring et al 2007). The reflectance spectra of the leaf material will be measured using a light spectrometer and methods established at RRes (Cook et al., 2013).

For all traits, correlation analyses will be used (UoY Bancroft) to identify those associated with variation in stress tolerance and candidate genes will be tested using radiation knock-out lines.

WP7. Understand the genetic and molecular basis of white rust tolerance

Milestones: **M7.1** Month 15 - *B. juncea* diversity panel analysed for tolerance to white rust (*Albugo candida*); **M7.2** Month 36 - Candidate genes for white rust tolerance validated; **M7.3** Month 36 - Candidate genes and pathways for *D. erucoides*–derived resistance to white rust in NRCPB *B. juncea* introgression lines identified.

Field trial will be conducted (in India) and assessed for diseases as appropriate to the location. Assessment of resistance to white rust by screening with core collection of white rust isolates will include the *B. juncea* genetic diversity panel and BC_2F_8 *B. juncea* introgression lines (derived at NRCPB through wide hybridization with *Diplotaxis erucoides*) and will be undertaken by NRCPB and IARI. Transcriptome sequencing of resistant and susceptible lines will be undertaken in replicated time course manner after pathogen inoculation and development under controlled epiphytotic conditions. NRCPB will purify RNA and commission Illumina sequencing (~20M x 150 base PE reads per sample). UoY Bancroft will analyse the mRNAseq data to identify genes differentially expressed. These will be used by NRCPB, with bioinformatics support from UoY Bancroft and in comparison with AT across the diversity panel, to identify candidate genes for resistance. These will be tested using radiation knock-out lines or genome editing, as appropriate.

WP8. Understand the genetic and molecular basis of stem rot tolerance

Milestones: **M8.1** Month 12 - Air sampling & analysis protocol for *Sclerotinia* spores established; **M8.2** Month 15 - *B. juncea* diversity panel analysed for tolerance to stem rot (*Sclerotinia sclerotiorum*); **M8.3** Month 24 – Diversity of pathogen population related to *B. juncea* in India assessed; **M8.4** Month 35 - Diversity of pathogenicity genes in *S. sclerotiorum* populations in India understood; **M8.5** Month 36 - *B. carinata* DH population developed and used for mapping stem rot resistance; **M8.6** Month 36 - Candidate genes for stem rot tolerance validated.

Diversity panel field trials will be assessed for stem rot. Dr. Navin Gupta (ICAR-National Research Centre on Plant Biotechnology) has germplasm with potential for *Sclerotinia* resistance. Transcriptome sequencing of resistant and susceptible *B. juncea* lines in a time course after pathogen inoculation and comparison with AT across the diversity panel will be used to identify candidate genes for resistance. UoH and Prof. Jon West (RRes) will collaborate with Dr. N. Gupta on RNAseq experiments and analyses with an emphasis on early time points after stem inoculation (Girard et al., 2017). Dr. N. Gupta's specifically aims to study transcriptomics and proteomics to identify candidate genes for stem rot tolerance. UoH will also use information of mapped QTL for *S. sclerotiorum* tolerance (Stotz et al., submitted) to exploit knowledge of disease resistance from other *Brassica* species (Wu et al., 2016).

UoH will cooperate with Dr. Pankaj Sharma (ICAR-DRMR), Dr. J. Kaur and Dr. N. Gupta to understand pathogen diversity and aggressiveness. UoH will receive lyophilised samples of infected leaves and spore tapes from the Indian partner to genotype *S. sclerotiorum* using known pathogenicity factors and neutral microsatellite markers (Mbengue et al., 2016, Attanayake et al., 2014). Dr. Sharma will study mycelial compatibility groups. The oxalate biosynthetic gene *oah1* encoding an oxaloacetate acetylhydrolase and the cutinase *SsCUT1* gene are of particular interest to UoH and RRes. Prof. J. West will identify *S. sclerotiorum* spores in the air in real time to develop an early warning system for fungicide applications and to identify coincidence of spore inoculum with flowering of lines in field experiments in India to differentiate resistance and disease escape.

A *B. carinata* diversity panel was screened previously by UDSC Pradhan and Kaur for stem rot resistance and lines showing significant tolerance were identified. These tolerant lines will be used to develop a DH mapping population (~200 plants). The population will be assessed for stem rot resistance in the field. QTL mapping of the resistance will identify targets for subsequent introgression into *B. juncea*.

WP9. Understand the genetic and molecular basis of black spot tolerance

Milestones: **M9.1** Month 15 - *B. juncea* diversity panel analysed for tolerance to black spot (*Alternaria brassicae*); **M9.2** Month 24 - Sequencing & annotation of the *A. brassicae* genome; **M9.3** Month 35 - Population genetic analysis of *A. brassicae*, including toxin genes, completed. **M9.4** Month 36 - Candidate genes and pathways for *D. erucoides*–derived resistance to *Alternaria* in NRCPB introgression lines identified; **M9.5** Month 36 - Candidate genes for black spot tolerance validated.

Diversity panel field trials will be assessed for black spot. Dr Bhattacharya/Dr Rao (ICAR-NRC on Plant Biotechnology), Dr Lakshman Prasad (IARI) and Prof Surinder Banga have identified *B. juncea-D.erucoides* introgression lines showing resistance to *Alternaria* and will be working on these lines particularly for identifying genes and pathways responsible for the resistance; development of markers associated with resistance will also be done. Transcriptome sequencing of resistant and susceptible *B. juncea* introgression lines and near isogenic at BC2F8, developed at ICAR-NRC on Plant Biotechnology, in a time course after pathogen inoculation and comparison with AT across the diversity panel will be used to identify candidate genes for resistance. These will be tested using radiation knock-out lines or genome editing or overexpression, as appropriate. In the absence of a published *A. brassicae* reference genome, UoH and partners (Dr. Lakshman Prasad, ICAR-IARI; Dr. P.D. Meena, ICAR-DRMR; Prof. J. West) will sequence, assemble *de novo* and annotate the pathogen genome using two or three selected isolates. This will support and help interpret RNAseq experiments of *A. brassicae*-infected *B. juncea*. Sequencing of the *A. brassicae* genome will also make available further analysis of pathogenicity genes and microsatellite markers. Of particular interest are host-selective toxins (Meena et al., 2017). Lyophilised samples of infected leaves and spore tapes will be analysed to determine pathogen diversity in different areas of India.

WP10. Understand the genetic and molecular basis of extreme virus resistance

Milestones: **M10.1** Month 24 – Diversity of TuMV in Indian *B. juncea* determined; **M10.2** Month 36 – Stability of TuMV TWBJ20 resistance to Indian TuMV isolates defined; **M10.3** Month 36 – Candidate TuMV resistance genes in *B. juncea* identified and validated by knock-outs and/or CRISPR/Cas9.

Turnip mosaic virus (TuMV) infects brassicas in India, including rapeseed mustard (Haq et al., 1994), however little is known about the diversity of the virus in India. At Warwick we have a large collection (>200 isolates) of TuMV from around the world and have characterised the biological, serological and molecular diversity of TuMV (Walsh & Jenner, 2002). Recently we identified a broad-spectrum, extreme resistance to TuMV in the *B. juncea* accession TWBJ20. Provisional characterisation and mapping indicates the resistance is effective against a range of TuMV pathotypes and controlled by two recessive genes on chromosomes A06 and A08 (Wang, 2016).

The diversity of TuMV in India will be analysed. *B. juncea* crops will be sampled and TuYV detected by enzyme-linked immunosorbent assay (ELISA). TuMV will be isolated from those plants where it is detected and the P3 gene of the virus, which has been shown to be responsible for the type and severity of symptoms (Jenner et al., 2003; Sanchez et al., 2015), will be sequenced and phylogenetic analyses performed. Isolates representing different phylogenetic groups will be used to challenge the TuMV-resistant *B. juncea* line TWBJ20. Challenged plants will be tested for the presence of TuMV using ELISA and RT-PCR to verify resistance and/or susceptibility. An existing BC_1 *B. juncea* population segregating for the TWBJ20 TuMV resistance will be exploited to fine map and identify candidate resistance genes. These will be investigated using radiation knock-out lines or genome editing, as appropriate to identify the actual resistance genes.

WP11. Understand the genetic and molecular basis of aphid and butterfly tolerance

Milestones: **M11.1** Month 15 - *B. juncea* diversity panel analysed for tolerance to mustard aphid and cabbage white butterfly; **M11.2** Month 36 - Candidate genes for insect tolerance validated.

Host plant preference of free-flying individuals of *B. juncea* pests will be tested with a focus on aphids (*Lipaphis erysimi*) and butterflies (*Pieris brassicae*). Field trials will be conducted in India and assessed for infestation by aphids and caterpillars (PAU; Sarwan, Banga). Larval survival and development of *P. brassicae,* colony growth of *L erysimi* and survival of *B. juncea* will be recorded. AT analysis of pest tolerance data will be undertaken (UoY Harper). First landings, time spent/plant, nymph production/oviposition preferences will be monitored with selected cultivars in laboratory bioassays (RRes Cook). Correlation analyses will be used to identify behavioral responses associated with leaf spectral reflectance, sucrose content, glucosinolate and wax profiles.

WP12. Understand the genetic and molecular basis of *Brassica*-broomrape interaction

Milestones: **M12.1** Month 12 - Method for culture of *Brassica-Orobanche* parasitism under glasshouse conditions; **M12.2** Month 24 - Identified germplasm for differential resistance to

Orobanche and root transcriptome of selected *Brassica* lines; **M12.3** Month 36 - Candidate genes and pathways for developing resistance to *Orobanche.*

Broomrape (*Orobanche spp*.) is a root parasite of Brassica crops and has emerged in last several years as the biggest threat to farmers in western and northern India. Little is known about the interaction so we propose one of the first studies. NRCPB Watts, Bhattacharya will develop methods for culture in the controlled conditions then screen the *B. juncea* diversity panel for variation in interaction. UoY Harper will undertake AT to identify candidate control genes. Root tissues from replicated trials of the most and least susceptible *B. juncea* accessions will be sampled by NRCPB, who will purify RNA and commission Illumina sequencing (~20M x 150 base PE reads per sample). UoY Bancroft will analyse the mRNAseq data to identify genes differentially expressed in roots. These will be used by NRCPB, with bioinformatics support from UoY Bancroft, to identify potential pathways for developing resistance.

WP13. Understand the genetic and molecular basis of temperature and drought tolerance

Milestones: **M13.1** Month 3 - Experimental parameters for *B. juncea* single and combined stress responses to drought and heat stress defined; **M13.2** Month 15 - *B. juncea* diversity panel analysed for combined heat and drought stress tolerance. **M13.3** Month 24 - Candidate drought tolerance genes identified by comparative genomics and network modelling in accessions with contrasting tolerance; **M13.4** Month 36 - Candidate genes for drought tolerance validated.

The *B. juncea* diversity panel will be assessed for drought and heat sensitivity based on final biomass parameters (biomass dry weights, pod numbers and seed yield). Segmented regression analysis on the soil dehydration profile will be used to calculate a drought sensitivity parameter (Ferguson et al 2017), and a proxy parameter for water use. We will validate the results by translating drought sensitivity parameters and combined stress responses of a subset of lines. This will include quantification of plant water use, whole-plant transpiration efficiency (total biomass per unit of water) and water productivity (seed biomass per unit of water). Photosynthetic snapshot measurements, δ13C on leaves/flowers that fully developed during the treatment conditions will be carried out to obtain an in-depth view of physiological status in relation to stress tolerance, water use and yield. The results of these experiments and AT across the diversity panel will be used to identify candidate genes for resistance, which will be tested using radiation knock-out lines or genome editing.

WP14. Understand the genetic and molecular basis of salinity tolerance

Milestones: **M14.1** Month 3 - Experimental design for high-throughput screening for root phenotyping during seedling stage established; **M14.2** Month 18 – Screening of five thousand accessions of *Brassica juncea* for root trait completed. **M14.3** Month 24 – Comparative transcriptomic analyses of tolerant and sensitive accession(s) completed; **M14.4** Month 36 – Functional validation of candidate genes in *Arabidopsis*/*Brassica juncea*.

Among several abiotic stresses, salinity is one of the important factor causing significant reduction in crop yield. In *Brassica juncea* varieties, the increase in mean salinity levels from 1.65 to 6.76 dS m-1 resulted in >70% reduction in overall seed yield (Chakraborty et al., 2015). Root is the only organ of plant that remains in direct contact to soil and information related to root phenotype in Indian mustard (*Brassica juncea*) is limited. Root architecture includes tap root and lateral roots and maintenance of optimum root architecture is the result of coordinated interaction between the genetic elements and environmental factors; growth and development of root is reported to be regulated by environmental stresses to optimize water and nutrient extraction (Quach, Truyen N., et al., 2014).

The *B. juncea* diversity panel along with 5000 accessions maintained at NBPGR, will be screened by state of art high-throughput platform under controlled condition at seedling stage to identify the potential candidate(s) showing robust root phenotype (root length, lateral roots) under saline conditions. Root tissues from tolerant lines will be sampled by NBPGR and NRCPB, who will purify RNA and commission Illumina sequencing (~20M x 150 base PE reads per sample). Transcript abundance will be quantified by UoY Bancroft and used for comparative transcriptomic analyses and, in combination with the gene expression associations identified in the diversity panel, to elucidate the underlying molecular network/genetic elements imparting salinity tolerance. Identified candidate gene(s) will be validated in *Arabidopsis*/*Brassica juncea* at ICAR-NRCPB using overexpression, radiation knock-out lines or genome editing, as appropriate.

WP15. Ensure effective data management to support pre-breeding

Milestones: **M15.1** Month 12 - Functional genotypes for *B. juncea* diversity panel available via the York Oilseed Rape Knowledgebase website; **M15.2** Month 36 - Molecular markers for marker-

assisted selection of studied traits made available via the York Oilseed Rape Knowledgebase website. **M15.3** Month 36 - Project trait data available via the Brassica Information Portal.

Lists of accessions being sequenced and key molecular data will be made publicly available via the York Oilseed Rape Knowledgebase site (http://www.yorknowledgebase.info/) by the UoY Bancroft group, with sequence data deposited in SRA. The *Brassica* Information Portal (https://bip.earlham.ac.uk/) will be expanded to accommodate *B. juncea* trait data via training of an Indian scientist hosted by EI Hall and the data made publicly available upon publication.

WP16. Ensure effective knowledge transfer

Milestones: **M16.1** Month 35 – UK industry and networking visit by Indian partners.

An important feature of the work plan is the training of Indian scientists, which will be accomplished largely by guided participation in the research. This will be supplemented by workshops and dissemination events. For example, the 2019 and 2021 UK Brassica Research Community meeting will be opened to the Indian partners to attend, present their work and network with UK *Brassica* researchers. During the 2021 visit they will be hosted by Elsoms Seeds Ltd to see and discuss the UK approach to breeding. Activities in India will focus on informing the supply chain of breeders (including the breeding arms of the participating institutions), farmers and agronomist.

WP17. Ensure effective project management

Milestones: **M17.1** Month 1- Initial consortium management meeting; **M17.2** Month 3 – germplasm exchange agreement signed; **M17.3** Month 35 - Final consortium management meeting.

Legal arrangements and 6-monthly project management meetings (by videoconference) and risk mitigation will be as described in the Partnership and Project Management Statement. A germplasm exchange agreement will be developed under the guidance of UDSC Pradhan, who has successfully exchanged germplasm with the UK (UoY Bancroft) previously.

Excellence and novelty

The programme exploits emerging draft genome sequences, adapting and applying to mustard rape (*B. juncea*) the latest genomic and bioinformatics approaches developed in the UK for oilseed rape (*B. napus*). For the first time, these are applied by a broad consortium of Indian experts, supported by the corresponding UK experts, to challenges facing commercial production.

d) The potential for improving oilseed productivity in India

As the largest oil crop in India, mustard rape has the greatest potential to increase oilseed production overall. It fits well in crop rotations systems and has well-established supply chains from breeders to consumers, making it the obvious oilseed choice for PORI. Stress-related losses are huge: yield losses due to *Alternaria* and white rust of 37% (Bal and Kumar, 2014). Small-scale farmers have limited resources to invest in crop protection, making genetic improvement an attractive approach.

References

Attanayake et al 2014 Heredity (Edinb) 113, 353-63. **Bal and Kumar 2014** Afr. J. Agric. Res. 9: 302-306. **Bancroft et al 2011** Nature Biotechnology 29:762-6. **Belfield et al 2012** Genome Research 22:1306-1315. **Chakraborty et al 2016** Journal of Plant Nutrition 39:7,1026-1037. **Cook et al 2013** Arthropod-Plant Interactions 7,249-258. **Döring & Chittka 2007** Arthropod-Plant Interactions 1,3-16. **Eigenbrode et al 2000** J Agri and Urban Ent 17,53-63. **Girard et al 2017** J. Exp. Bot. **Haq et al 1994** Journal of Phytopathology 140,55-64. **Harper et al 2012** Nature Biotechnology 30:798-802. **He et al 2015** Data in Brief 4:357-362. **He et al 2016** Plant Biotechnology Journal, DOI: 10.1111/pbi.12657. **Hervé et al 2014** J chemical ecology 40,1220-1231. **Jenner et al 2003** Molecular Plant-Microbe Interactions 16,777-784. **Kumar et al 2009** (http://www.australianoilseeds.com/__data/assets/pdf_file/0009/6849/31_Canola_cultivation_in_India_scenario_and_future_strategy.pdf). **Mbengue et al 2016** Front Plant Sci 7,422. **Meena et al 2017** Front Microbiol 8,1451. **Miller et al 2016** BMC Genomics 17:500. **Ni et al 2013** Plant Pathol 63,174–184. **Quach et al 2014** PLoS ONE 9(1):e84886. **Sánchez et al 2015** Molecular Pant-Microbe Interactions 28,1304-1315. **Sollars et al 2017** Nature 541:212-216. **Trick et al 2009** Plant Biotechnology J. 7:334-346. **Walsh & Jenner 2002** Molecular Plant Pathology 3,289-300. **Wang 2016** PhD Thesis, University of Warwick, 147pp. **Wu et al 2016** Front Plant Sci 7,1418. **Yang et al 2016** Nature Genetics 10:1225-1232. **Zamir 2001** Nature reviews Genetics 2:983-989.