Biological Journal of the Linnean Society, 2005, **84**, 637–641. With 2 figures

The genus Mus as a model for evolutionary studies
Edited by J. Britton-Davidian and J. B. Searle

Positional cloning of the Hybrid sterility 1 gene: fine genetic mapping and evaluation of two candidate genes

ZDENĚK TRACHTULEC[1]*, ONDREJ MIHOLA[1], CESTMÍR VLCEK[1], HEINZ HIMMELBAUER[2], VÁCLAV PAČES[1] and JIŘÍ FOREJT[1]

[1]*Centre of Integrated Genomics, Institute of Molecular Genetics, Academy Sciences of the Czech Republic, Videnská 1083, CZ-14220 Prague 4, Czech Republic*
[2]*Max Planck Institute for Molecular Genetics, Ihnestrasse 73, 14195 Berlin, Germany*

Received 3 September 2003; accepted for publication 7 October 2004

The Hybrid sterility 1 (*Hst1*) gene affects fertility of male hybrids between certain laboratory strains (such as C57BL/10) and some *Mus musculus musculus* mice by causing a breakdown of spermatogenesis at the stage of primary spermatocytes. In the process of positional cloning of the *Hst1* gene, we generated a contig of bacterial artificial chromosomes and subsequently a low coverage sequence of the candidate region of the 129S1/SvImJ strain. Development of new genetic markers allowed us to narrow the *Hst1* region from 580 to 360 kb. The products of two genes from this region, TATA-binding protein (*Tbp*) and proteasome subunit beta 1 (*Psmb1*), accumulate during spermatogenesis. These proteins have been described previously as having conserved C-terminal sequences and species-specific N-termini. We evaluated the candidacy of these genes for *Hst1* by allelic sequencing and by real-time semiquantitative reverse-transcription PCR of testicular mRNAs. *Biological Journal of the Linnean Society*, 2005, **84**, 637–641.

ADDITIONAL KEYWORDS: house mouse – *Mus musculus* – speciation.

INTRODUCTION

Hybrid sterility prevents gene flow between related species, and it is therefore important in the process of speciation (for a review see Coyne, 1992). To understand speciation, it is important to identify the genes responsible for the sterility of hybrids. The Hybrid sterility 1 (*Hst1*) gene on mouse chromosome 17 was the first such gene mapped in mammals (Forejt & Ivanyi, 1974). The gene affects fertility of male hybrids between certain laboratory strains (such as C57BL/10SnPh, henceforth B10) and some *Mus musculus musculus* mice (e.g. of the PWD/Ph strain) by causing a breakdown of spermatogenesis at the stage of primary spermatocytes. Other laboratory strains, e.g. C3H/DiSnPh (hereafter C3H), produce fertile males in these intersubspecific crosses (for a review see Forejt, 1996). The laboratory strains B10 and C3H are predominantly of *M. m. domesticus* origin (Bonhomme *et al*., 1987). The difference between the B10 and C3H strains is due to a single gene, *Hst1*, which was mapped using the backcross (B10-*T* × C3H)*T* × B10 (Forejt *et al*., 1991; Gregorova *et al*., 1996). In the process of positional cloning of *Hst1*, the gene was localized in a 2-Mb region of chromosome 17 covered with yeast artificial chromosomes (YACs; Trachtulec *et al*., 1994). Genes were also identified and mapped to this region (Trachtulec *et al*., 1997a, b; Trachtulec & Forejt, 1999).

The YAC clones are difficult to handle and process. To obtain clones more suitable for transgenesis, knockout and replacement experiments, and for refinement of the genetic and physical maps down to

*Corresponding author. E-mail: trachtul@biomed.cas.cz

the nucleotide sequence, bacterial artificial chromosomes (BACs) were chosen. A contig of BAC clones and a low coverage sequence of the candidate region of the 129S1/SvImJ strain were generated. New genetic markers were then developed and the *Hst1* region was characterized more precisely. The candidacy of two genes was tested by allelic sequencing and by quantification of testicular transcripts.

MATERIAL AND METHODS

Six nylon filters containing a gridded BAC genomic library (ResGen) from the mouse CJ7 embryonic stem cell line DNA inserted into the *Hin*dIII site of the pBeloBAC11 vector were used. The cell line most closely resembles the mouse strain 129S1/SvImJ. Probes were labelled by random decamer priming and hybridized to the filters in 50% formamide, 4× SSPE, 1% sodium dodecyl sulphate (SDS), 10× Denhardt's solution, and 0.05 mg/mL carp sperm DNA at 42 °C overnight. The filters were then washed at high stringency (0.4× SSPE, 0.2% SDS at 67 °C) and autoradiographed. The primer sequences and PCR conditions used for rescreening were as in Trachtulec & Forejt (1999). The BAC DNAs were prepared on Qiagen columns, digested by one or two restriction endonucleases and resolved by pulse-field gel electrophoresis (CHEF DRIII, Bio-Rad) on 1% agarose. The gels were blotted, hybridized and washed as indicated above. The probes were PCR fragments of the ten loci used for rescreening (see Results) and clones containing the *Psmb1* and *Dll1* cDNAs. The BAC DNAs were subcloned into mp13mp18 or pUC18 vectors after partial digestion with *Tsp*509 I (for BACs 356I8 and 120O15) or complete digestion with *Hin*dIII, *Bgl*II or *Bam*HI (120O15). The DNAs of the subclones were sequenced by the CEQ DTCS-Quick Start Kit using a capillary sequencer CEQ2000XL (BeckmanCoulter). The reads were aligned with the GeneSkipper software (EMBL).

Allelic sequencing of the *Tbp* and *Psmb1* mRNAs was performed by PCR on testicular cDNA libraries. The PCR primers and libraries are given in Trachtulec *et al*. (1997a).

Quantification of mRNAs was performed using a real-time PCR system (LightCycler, Roche). Total RNA was prepared using Trizol (Invitrogen). Reverse transcription (RT) was performed with MuMLV RTase (Invitrogen). An aliquot corresponding to 50 ng RNA was added to each reaction with FastStart DNA Master SYBR Green I kit and cycled in LightCycler. As controls, RT reactions without reverse transcriptase were utilized. The primers (5′ to 3′) were: for *Tbp*, ATCCCAAGCGATTTGC and GCTCCCCACCAT GTTC (annealing at 61 °C, product 200 bp); for *Psmb1*, CGCTTCTTCCCTTACTATGTT and TTTC CTCCCTGATGCCCTCTT (60 °C, 320 bp); and for *Gapdh*, AATGTGTCCGTCGTGG and CTGGTCCT CAGTGTAGCC (61 °C, 100 bp). The data were analysed using LightCycler Software version 3.5.3 (Roche).

RESULTS AND DISCUSSION

To obtain BAC clones covering the *Hst1* region, the 129S1/SvImJ mouse BAC library, spotted on nylon filters, was hybridized to the mixture of probes for loci *D17Aus17* and *Zfp91-rs1*. The probes have been characterized previously (Trachtulec *et al*., 1997b). Prescreening of the clones was done by PCR with the *Zfp91-rs1* and *D17Aus17* primers on bacterial colonies. Eighteen positives were isolated. These clones included two pairs from neighbouring wells, which were likely to be identical. To order clones on the map, another round of screening by eight additional pairs of PCR primers was performed. The PCR assays identified loci *Tbp*, *Pdcd2*, *D17Ph4e*, *D17Ph6*, *D17Ph7* and *D17Mia14–Mia16*. Six BAC clones: 120O15, 244F22, 131P12, 357 N17, 237M12, and 356I8, were chosen for restriction mapping with six endonucleases (Fig. 1). The resulting map corresponded well to the previously established YAC and long-range genomic DNA maps (Trachtulec *et al*., 1997b). This result indicated a lack of gross rearrangements between the mouse strains C57BL/6, C3H and 129 in the *Hst1* region. The order of the loci determined previously (Trachtulec & Forejt, 1999) was as in this study. The minimal *Hst1* co-segregating region was previously mapped to the region encompassed by the loci *D17Ph4e* and *Zfp91-rs1* (Trachtulec *et al*., 1997b). As the BAC clones 120O15 and 356I8 extend to previously non-characterized regions, they were selected to shotgun subcloning and random sequencing. Eighty sequencing reads were obtained from the 356I8 clone and 200 reads from the 120O15 clone. The sequences containing putative polymorphisms with the published C57BL/6J sequence (Waterston *et al*., 2002) were deposited in the EMBL/GenBank/DDBJ databases under accession numbers AY337854–AY337862. Eleven pairs of primers flanking microsatellites were selected based on the sequence. The oligonucleotides were then utilized for PCR to test the parental strains of our backcross (B10 and C3H; Gregorova *et al*., 1996) for a microsatellite length difference on 4% agarose. Four of the microsatellites were polymorphic and the corresponding primers were therefore used to map the *Hst1* gene on our backcross. The list of the primers is given in Table 1. One of the loci co-segregated with *Hst1*, whereas three others were separated from it by two recombinants (0.15 cM) each. Previously, the *Hst1* gene was located by the combination of physical and genetic mapping to a region 580 kb in length (YAC end markers *D17Mia7* and *D17Ph1*; Gregorova *et al.*, 1996). The new closest segregating markers (DB3 and CG53) map 360 kb

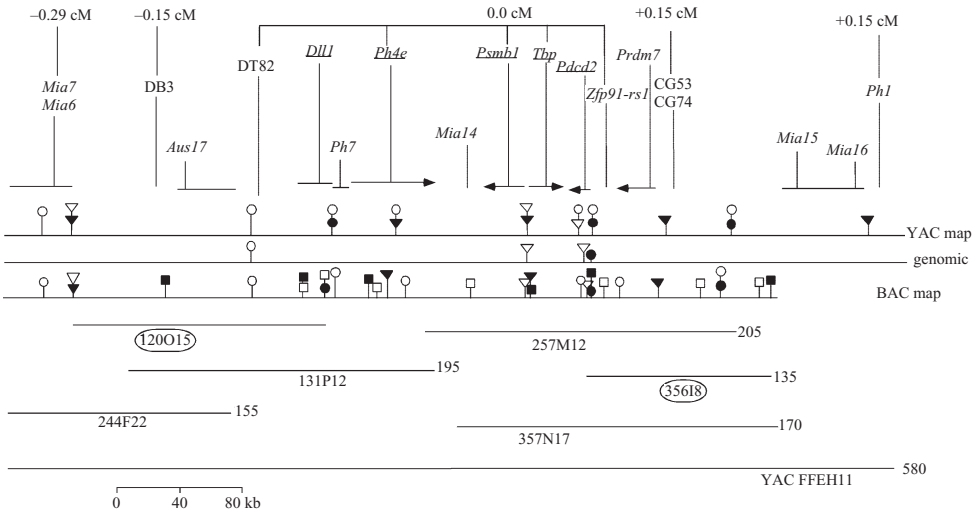

Figure 1. Maps of the *Hst1* region. The genetic distances proximal (minus) and distal (plus) to the *Hst1* gene are given at the top. The six candidate genes are underlined. The *D17* prefix is omitted from some symbols. The *Mia* and *Ph* markers, the YAC map and total genomic PFGE DNA map have been published previously (Trachtulec *et al*., 1994, 1997a, b). The BAC consensus map was constructed by aligning maps of six independent BAC clones. The BAC clones are shown as horizontal lines with the clone numbers below and size (in kb) on the right. The two partially sequenced BACs are circled. The FFEH11 YAC clone containing *Hst1* is at the bottom. The enzymes used for the total genomic map were: *Not*I, filled circles; *Mlu*I, open circles; *Nru*I, open triangles; for the YAC map, *Ecl*XI (cuts also at the *Not*I sites), filled triangles, was also utilized. Two additional nucleases, *Bss*HII (closed rectangles) and *Sal*I (open rectangles), were used for the generation of the BAC map.

Table 1. Fine genetic mapping of the *Hst1* gene with microsatellites

Sequence name	BAC clone ID number	Primer sequences (5′–3′)	Polymorphism*	Chromosome 17 coordinates (kb)	Distance to *Hst1*
CG74	356I8	TTCCTAACACAGAGTAAATCAT, GACTGCACAGGTTAGTTACAGA	B10 > C3H	14 117	+0.15 cM
CG53	356I8	TCTGAAGAGACTACATTTGATG, GGTAACCTTGATAGAGTAGATTT	B10 > C3H	14 118	+0.15 cM
DT82	120O15	TTTTCTTGCTGTGACTGTTTTAT, CGGCCATACTCATCTCTACA	B10 > C3H	13 843	0.0 cM
DB3	120O15	GGATCCTGGTTCTCATCTCTC, AGAAAGACAAAAAGCAAGCATAC	B10 < C3H	13 759	–0.15 cM

*The relative length of the alleles.
The coordinates are according to Ensembl assembly 3 (release 13.30.1, May 2003).

apart, between 13 759 kb and 14 118 kb on the physical map of chromosome 17; their distance was estimated by BLASTing the sequences of the loci against the C57BL/6J draft mouse sequence (Ensembl v. 16.30.1) and extracting the coordinates of the hits (Table 1). The minimal region of the *Hst1* gene, defined by the outermost co-segregating markers, was extended from about 130 kb (*Zfp91-rs1* to *D17Ph4e*) to 200 kb (*Zfp91-rs1* to DT82).

The *Hst1* region now encompasses six candidate genes: Delta-like 1 (*Dll1*), *D17Ph4e* or *4932442K08Rik* with 5′3′-exonuclease domain and ATP/GTP-binding site motif A (P-loop), proteasome subunit beta 1 (*Psmb1*), TATA-binding protein (*Tbp*), programmed cell death 2 (*Pdcd2*) and a homologue of the human PR-domain 7 gene (*Prdm7*). There are also at least two pseudogenes, *Zfp91-rs1* and *Odc-rs15*. The product of the candidate gene *Tbp* (TBP) is required to initiate transcription with all three types of RNA polymerases (Hernandez, 1993). The TBP accumulates during spermatogenesis (Schmidt & Schibler, 1997). The PSMB1 protein participates in the function both

of the 19S and of the 26S proteasome (also called proteosome, prosome, or multicatalytic endopeptidase complex), the major non-lysosomal ATP-dependent protein degradation system (Goldberg, 1995). The proteasomes are required for replacement of histones by transition proteins and protamines and for remodelling of sperm tails during spermiogenesis (Tipler *et al*., 1997; Mochida, Tres & Kierszenbaum, 2000). The TBP and PSMB1 proteins were also described as having conserved C-terminal sequences and species-specific N-termini (Hernandez, 1993; Savioz, Houghton & Davies, 1995). Therefore, these two genes were chosen for detailed analysis by allelic sequencing and mRNA quantification. The testicular transcripts of *Tbp* and *Psmb1* were isolated and sequenced from both the B10 and the C3H strains. The sequences were deposited with the EMBL/GenBank/DDBJ under accession numbers U63933 (*Tbp*) and U64452 (*Psmb1*). One silent (synonymous) nucleotide substitution was found in the coding region of both the *Tbp* and the *Psmb1* gene. In addition, differences were detected in the untranslated regions. Thus, the predicted proteins are identical. However, the sequence differences could affect the stability of mRNA, and, moreover, there could be differences in DNA regulatory sequences of these genes, not investigated for polymorphisms. A previous study has shown that significant differences between the frequency of pachytene spermatocytes in testis of prospectively fertile and sterile hybrids can be detected on day 15 (10% vs. 2%, Forejt, 1981), indicating that the *Hst1* gene could be differentially expressed at this stage of spermatogenesis or earlier. Therefore, the levels of transcripts of *Tbp* and *Psmb1* were analysed in 15-day-old mouse testis of prospectively fertile and sterile hybrids and controls by real-time semiquantitative reverse-transcription polymerase chain reaction (QRT-PCR). The primers were chosen to detect the region common to all known transcript alternatives. Both genes displayed similar transcript levels in the 15-day-old mouse testis RNA samples (Fig. 2). The adult testis mRNA levels of the *Tbp* are higher in the fertile than in the sterile mice, in agreement with the elevated expression of this gene during spermiogenesis (Schmidt & Schibler, 1997). These stages are largely missing in the sterile hybrids (Forejt & Ivanyi, 1974). The allelic sequences and mRNA levels suggest that neither *Tbp* nor *Psmb1* is the *Hst1* gene.

In summary, the *Hst1* region was reduced to 360 kb by physical and genetic mapping, the BAC clones suitable for genetic manipulation of the region were obtained and two genes were excluded as candidates for *Hst1*. We are proceeding with allelic sequencing and mRNA quantification of the other genes and conserved stretches from the candidate region. The

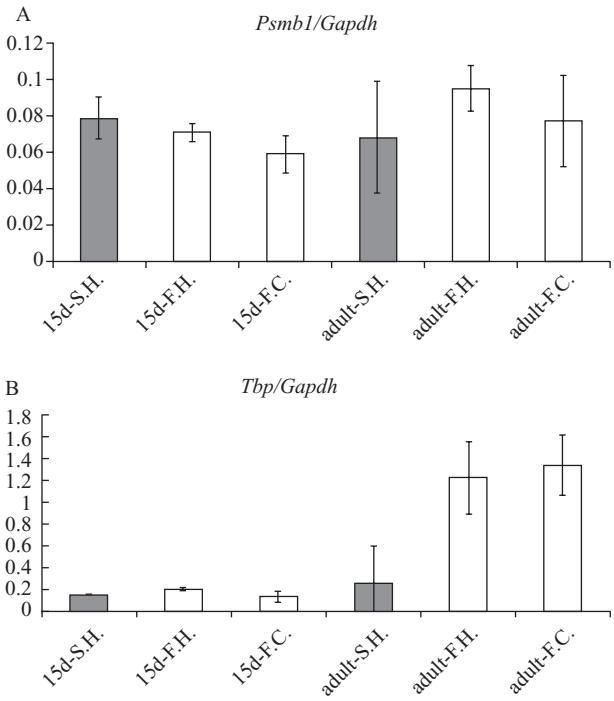

Figure 2. Semiquantitative analysis of testicular mRNA levels of the *Psmb1* (A) and *Tbp* (B) genes relative to the *Gapdh* gene. The relative expression levels (arbitrary units) for the (prospectively) sterile hybrid testis are indicated in dark grey. 15d, 15-day-old testis; adult, adult testis; S.H., sterile hybrid [genotype F_1(PWD × B10)]; F.H., fertile hybrid (reciprocal F_1 crosses of C3H and PWD); F.C., fertile controls: strains of *Mus musculus domesticus* origin (B10, C3H), and of *M. m. musculus* origin (PWD). The experiments were repeated at least twice on 2–6 independently collected samples of each group and with different RT reactions.

allelic differences detected in the *Tbp* and *Psmb1* genes are being used for a haplotype analysis in laboratory and wild mice. In nature, *M. m. musculus* and *M. m. domesticus* meet in a narrow hybrid zone (for a review see Boursot *et al*., 1993; see also Božíková *et al.*, 2005; Dod *et al.*, 2005; Raufaste *et al*., 2005, all this issue). Our preliminary results suggest that the haplotypes carrying the *Hst1* alleles of the laboratory strains also occur in wild *M. m. domesticus* populations, which makes the *Hst1* a candidate speciation gene and not an artefact which arose during the production of laboratory strains. However, some *M. m. musculus* mice can produce fertile hybrids with *Hst1*^{s}-carrying laboratory strains (Forejt & Ivanyi, 1974; Vyskočilová *et al*., 2005, this issue). Therefore, only a proportion of F_1 male hybrids are expected to be sterile (females being always fertile) in nature.

ACKNOWLEDGEMENTS

We thank to M. Landikova and I. Opltova for technical assistance and S. Takacova for careful reading of the manuscript. We were supported by grants from the Czech Ministry of Education (No. LN00A079) and from the Grant Agency of the Czech Republic (204/98/KO15 and 204/01/0997). J.F. is an International Research Scholar of the Howard Hughes Medical Institute.

REFERENCES

Bonhomme F, Guénet J-L, Dod B, Moriwaki K, Bulfield G. 1987. The polyphyletic origin of laboratory inbred mice and their rate of evolution. *Biological Journal of the Linnean Society* **30:** 51–58.

Boursot P, Auffray J-C, Britton-Davidian J, Bonhomme F. 1993. The evolution of house mice. *Annual Review of Ecology and Systematics* **24:** 119–152.

Božíková E, Munclinger P, Teeter KC, Tucker PK, Macholán M, Piálek J. 2005. Mitochondrial DNA in the hybrid zone between *Mus musculus musculus* and *Mus musculus domesticus*: a comparison of two transects. *Biological Journal of the Linnean Society* **84:** 363–378.

Coyne JA. 1992. Genetics and speciation. *Nature* **355:** 511–515.

Dod B, Smadja C, Karn RC, Boursot P. 2005. Testing for selection on the androgen-binding protein in the Danish mouse hybrid zone. *Biological Journal of the Linnean Society* **84:** 447–459.

Forejt J. 1981. Hybrid sterility gene located in the *T*/*t- H-2* supergene on chromosome 17. In: Reisfeld S, Ferone RA, eds. *Current trends in histocompatibility*. New York: Plenum Press, 103–131.

Forejt J. 1996. Hybrid sterility in the mouse. *Trends in Genetics* **12:** 412–417.

Forejt J, Ivanyi P. 1974. Genetic studies on male sterility of hybrids between laboratory and wild mice (*Mus musculus* L.). *Genetical Research* **24:** 189–206.

Forejt J, Vincek V, Klein J, Lehrach H, Loudova-Mickova M. 1991. Genetic mapping of the *t*-complex region on mouse chromosome 17 including the Hybrid sterility-1 gene. *Mammalian Genome* **1:** 84–91.

Goldberg AL. 1995. Functions of the proteasome: the lysis at the end of the tunnel. *Science* **268:** 522–523.

Gregorova S, Mnukova-Fajdelova M, Trachtulec Z, Capkova J, Loudova M, Hoglund M, Hamvas R, Lehrach H, Vincek V, Klein J, Forejt J. 1996. Sub-milli-Morgan map of the proximal part of mouse Chromosome 17 including the hybrid sterility 1 gene. *Mammalian Genome* **7:** 107–113.

Hernandez N. 1993. TBP, a universal eukaryotic transcription factor? *Genes and Development* **7:** 1291–1308.

Mochida K, Tres LL, Kierszenbaum AL. 2000. Structural features of the 26S proteasome complex isolated from rat testis and sperm tail. *Molecular Reproduction and Development* **57:** 176–184.

Raufaste N, Orth A, Belkhir K, Senet D, Smadja C, Baird SJE, Bonhomme F, Dod B, Boursot P. 2005. Inferences of selection and migration in the Danish house mouse hybrid zone. *Biological Journal of the Linnean Society* **84:** 593–616.

Savioz A, Houghton I, Davies RW. 1995. Cloning and sequencing of a murine cDNA encoding the proteasome component C5. *DNA Sequence* **5:** 307–309.

Schmidt EE, Schibler U. 1997. Developmental testis-specific regulation of mRNA levels and mRNA translational efficiencies for TATA-binding protein mRNA isoforms. *Developmental Biology* **184:** 138–149.

Tipler CP, Hutchon SP, Hendil K, Tanaka K, Fishel S, Mayer RJ. 1997. Purification and characterization of 26S proteasomes from human and mouse spermatozoa. *Molecular Human Reproduction* **3:** 1053–1060.

Trachtulec Z, Forejt J. 1999. Transcription and RNA processing of mammalian genes in *Saccharomyces cerevisiae*. *Nucleic Acids Research* **27:** 526–531.

Trachtulec Z, Hamvas RM, Forejt J, Lehrach HR, Vincek V, Klein J. 1997a. Linkage of TATA-binding protein and proteasome subunit C5 genes in mice and humans reveals synteny conserved between mammals and invertebrates. *Genomics* **44:** 1–7.

Trachtulec Z, Mnukova-Fajdelova M, Hamvas RM, Gregorova S, Mayer WE, Lehrach HR, Vincek V, Forejt J, Klein J. 1997b. Isolation of candidate hybrid sterility 1 genes by cDNA selection in a 1.1 megabase pair region on mouse chromosome 17. *Mammalian Genome* **8:** 312–316.

Trachtulec Z, Vincek V, Hamvas RM, Forejt J, Lehrach H, Klein J. 1994. Physical map of mouse chromosome 17 in the region relevant for positional cloning of the Hybrid sterility 1 gene. *Genomics* **23:** 132–137.

Vyskočilová M, Trachtulec Z, Forejt J, Piálek J. 2005. Does geography matter in hybrid sterility in house mice? *Biological Journal of the Linnean Society* **84:** 663–674.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E *et al.* 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420:** 520–562.